IRSTI 14.01.73 10.51889/2960-1649.2024.60.3.015

Z.K. JUMABAYEVA[1], A.S. SHILIBEKOVA[1], N.M.A. DIETEREN[2], A.S. JANDAROVA[1], N.T. SHORA[1]*

[1]*Nazarbayev Intellectual schools (Astana, Kazakhstan)*
[2] *Centraal Instituut voor Toetsontwikkeling (Arnhem, Netherlands)*
**e-mail: jumabayeva_zh@cpi.nis.edu.kz*

NAZARBAYEV INTELLECTUAL SCHOOLS' UPDATED STUDENT SELECTION SYSTEM: ANALYSIS OF THE EFFECTIVENESS AND VALIDITY

Abstract

This paper presents the validation research conducted in 2020 for a testing system designed to identify prospective students for a network of selective STEM schools. The initial validation, carried out in 2015, established that the assessment system was fit for purpose, effectively selecting suitable candidates. Since then, the selection process has undergone modifications, prompting the need for further validation. This study evaluates the effectiveness and predictive validity of the updated selection system, focusing on students in grade 7. Data were collected from a single cohort during the 2019-2020 academic year, focusing on their assessment results. The findings indicate that the selection test successfully predicts academic achievement in mathematics and quantitative reasoning, with moderate predictive accuracy in languages. Despite changes to the selection test since 2019, the overall predictive validity of the subtests remains strong. The paper emphasizes the practical significance of these findings for educational institutions, offering insights to improve admission processes. By refining the selection system, educational organizations can better identify and support talented students, ultimately improving educational outcomes. The study also contributes to the ongoing improvement of assessment practices, ensuring that selection procedures remain effective in identifying students capable of thriving in rigorous academic environments.

Keywords: validation research; selection test; predictive validity; quantitative reasoning; intellectual school

Introduction. Nazarbayev Intellectual schools (NIS) constitute a system of STEM-based (science, technology, engineering, and mathematics) educational institutions that specifically target exceptional and motivated students during their selection process. STEM education, initially conceived as a meta-discipline that integrated knowledge from various fields, has now become a focal point for many educators and policymakers, emphasizing the integration of STEM disciplines in schools (Chomphuphra et al., 2019). Admission to NIS is based on competitive examinations. As of 2013, the selection process included tests in mathematics, languages (Kazakh, Russian and English), and the ability test (quantitative reasoning and spatial reasoning).

Since 2019, the NIS selection test has undergone several modifications to align with the updated curriculum content taught to students from 6th grade onwards. As a result, the student selection test now follows a new format to accommodate these changes and includes tests in mathematics, quantitative reasoning, and languages (Kazakh, Russian and English) (Center for Pedagogical Measurements, 2023). As can be seen, the composition of the ability test changed, leaving only the quantitative reasoning subtest as part of the selection test procedure. The total score scale for selection test was also extended from 1000 to 1300. Changes also affected the mathematics subtest booklet, reflecting the updated curriculum.

This paper is focused on 2020 validation research conducted to evaluate the effectiveness of the testing system used in the selection of prospective students for enrollment at NIS. For the first time in 2015, validation research was carried out to verify the appropriateness of the assessment systems and to ensure that the selection process accurately identified the suitable candidates for admission to NIS.

In the paper we will particularly explore the effectiveness and predictive validity of the student selection system for 2,925 students in grade 7 from 19 schools. The predictive validity of 2019 selection test was analysed for the first cohort that took this test in the updated version.

As quantitative data, the results of mathematics, quantitative reasoning, and languages (Kazakh, Russian and English) subtests were extensively used. These results were correlated with students' scores for each grade 7 subject at the end of each term and their monitoring assessment results (Nazarbayev Intellectual schools, 2023).

The key question addressed in this study is whether modifications to the composition of the selection test and its subtests have led to significant changes in their predictive validity. To explore this question, the analysis is divided into two sub-questions aimed at assessing the effectiveness of the test in identifying appropriate candidates. These sub-questions are: (1) have suitable students been selected, and (2) have students been selected in the right way. The following sections will examine these issues in greater depth to determine if notable changes have occurred in the predictive validity of the test and its components following the recent modifications.

Composition of the selection test. It is important to highlight that the administration of the selection test in 2019 differed from the previous administrations since 2013. The table 1 provides an overview of these differences. These notable changes provided a compelling rationale for conducting fresh validation research following the 2019 selection test administration.

Table 1. *Composition and structure of the NIS selection test*

	2013-2018	2019
Day 1	Mathematics *(60 minutes)* Kazakh/Russian as 1st language *(40 minutes)* Kazakh/Russian as 2nd *language (40 minutes)* English language *(40 minutes)*	Mathematics *(60 minutes)* Quantitative Reasoning *(30 minutes)*
Day 2	*77 minutes* Quantitative Reasoning Spatial Reasoning	*120 minutes* Kazakh language (1st or 2nd) Russian language (1st or 2nd) English language

As it can be seen from the table, the composition of the subject's test remained the same, but the languages subtests were shifted from day 1 to day 2. Total testing time for these subjects remained the same, but instead of having three separate 40-minute sessions for languages subtests, they were combined into one longer session that lasted 120 minutes. This single session included all the language test questions in a single booklet.

As mentioned above, the composition of the CTY abilities test changed, leaving only the quantitative reasoning subtest as part of the selection test procedure.

Furthermore, there were some new topics introduced in the mathematics subtest booklet, according to the revised test matrix that was based on the updated curriculum.

Finally, the scores from five subtests were added together to calculate a total score, which had a maximum possible value of 1300. This combined score was then used to create a ranking for all candidates per school (city) and per language stream (Kazakh or Russian). For the subtests Mathematics and QR there is also a minimum passing score applicable, what means that candidates who do not reach one or both pass scores will not be selected, whatever their overall score might be.

Over time, the student selection process has evolved to include two comprehensive paper-based tests: the STEM test (mathematics, quantitative reasoning, natural science sections) and language test (Kazakh, Russian and English) (Nazarbayev Intellectual schools, 2023). However, the 2019 test does not cover the 2022 natural sciences subtest.

Materials and Methods. To validate the selection test, we first analyzed the relation between school success of the students that have entered grade 7 in NIS (results monitoring tests and the summative scores on terms), the results

on the selection test and background variables such as language, school and region. The method we use is LISREL (LInear Structural RELations).

There are two research sub-questions:

1: Have the suitable students be selected?

Did the selection test effectively support the selection of the students that have the highest chance of being successful in the NIS educational career? This research question can be answered by calculating the degree of correlation between individual learners' performances in the selection test (total and per subtest) and the performances of the same learners in all subjects and terms in grade 7.

2: Are the students selected in the right way?

Is the composition and the structure of the test appropriate for the aim of the assessment? In this question we want to relate elements of the design of the test, like number of subtests, number of items in relation to assessment time and assessment objectives, to the purpose: selecting for grade 7 the most gifted learners who have the highest probability of achieving the final qualification of NIS at grade 12.

For this research we used data of one cohort: the grade 7 cohort of 2019-2020 academic year, that participated in the selection test in 2019. The goal is to predict successfulness of selection and the differences between certain subgroups/ subdimensions. The predictive validity of the selection test and each of the subtests must be explored and described. The limitation of this research approach is that we do not have data about students' performance- over the whole NIS career. For this research we assume that students who are most successful in achieving high marks/scores in grade 7 will also be the most successful students when doing the final qualification at the end of NIS grade 12.

Correlation research of the results of selection test 2019 was broken down into the subtests performance, with separate attention for the QR subtest. The QR subtest underwent substantial changes in the content and underlying test matrix. And 2019 was the first time that the subscore on the QR-test was fully included in the overall scoring of selection test results.

Data Methods. The data that were based on the assessment results of one cohort of NIS students: the selected grade 7 candidates from the 2019 selection test cycle. The available data contained the following parameters (See Table 2)**.**

Table 2: *Data for the 2020 validation research of NIS selection test*

Field name	Data type	Constraint	Description	Notes
iin	Text	Primary Key	Student national ID number	12-digit Individual Identification Number
school_id	Number	Not null	School (NIS) ID	1-2, 4-21
gender	Number	Not null	Student gender	1 – male, 2 – female
lang	Number	Not null	Student language of study	1 – kazakh, 2 – russian
locality	Text	Not null	Student urban-rural classification based on previously attended school	1 – urban, 2 – rural
CTY group	Text	Not null	Students' abilities to study natural and mathematical sciences	Levels: I – very high, II – high, III – average, IV – sufficient
ko_math	Number	Not null	Selection test Mathematics results	Max – 400

ko_qr	Number	Not null	Selection test Quantititave Reasoning results	Max – 300
ko_d1_score	Number	Not null	Selection test Day I results	Max – 700
ko_kz	Number	Not null	Selection test Kazakh language results	Max – 200
ko_ru	Number	Not null	Selection test Russian language results	Max – 200
ko_en	Number	Not null	Selection test English language results	Max – 200
ko_ d2_score	Number	Not null	Selection test Day 2 results	Max – 600
ko_total	Number	Not null	Selection test Final results	Max – 1300
SOR_SUBJECT_ TERM	Number	-	Summative score per section	Terms 1–4 (I–IV). Max – 50
SOCH_SUBJECT_ TERM	Number	-	Summative score per term	Terms 1–4 (I–IV). Max – 50
ML919_SUBJECT	Number	-	Monitoring Language results (September 2019)	Max – 50
level_SUBJECT	Text	-	Monitoring Language results classification (September 2019)	Levels: Base, Beginner, Good, Advanced
1N_M19	Number	-	Monitoring Mathematics results (September 2019)	1N – Numbers. Max – 50
2A_M19	Number	-	Monitoring Mathematics results (September 2019)	2A – Algebra. Max – 50
3G_M19	Number	-	Monitoring Mathematics results (September 2019)	3G – Geometry. Max – 50
4S_M19	Number	-	Monitoring Mathematics results (September 2019)	4S – Statistics. Max – 50
5M_M19	Number	-	Monitoring Mathematics results (September 2019)	5M – Mathematical modeling. Max – 50
1N_M20	Number	-	Monitoring Mathematics results (January 2020)	1N – Numbers. Max – 50
2A_M20	Number	-	Monitoring Mathematics results (January 2020)	2A – Algebra. Max – 50
3G_M20	Number	-	Monitoring Mathematics results (January 2020)	3G – Geometry. Max – 50
4S_M20	Number	-	Monitoring Mathematics results (January 2020)	4S – Statistics. Max – 50
5M_M20	Number	-	Monitoring Mathematics results (January 2020)	5M – Mathematical modeling. Max – 50

After careful exploration of the data provided, the research team decided that all data could be used, except for the CTY group level indication. This parameter was the only assessment-results-based parameter that was not expressed in continuous scale measure units, but rather in four distinct group labels. Furthermore, the group classification was not based on a given formula or algorithm but on a combined judgment of two separate parameters: the QR score in the 2019 high stakes selection test and a Spatial reasoning (SR) score as result of a low stakes in-class administration of this subtest during the school year 2019-20.

Data with background variables of the students and data with social-emotional indicators, were not used in this 2020 research. The main goal of this validation research is to determine if the changes made to the selection test have significantly impacted its effectiveness and predictive validity. The 'predictive quality' of the test items and subtests should not be differential for specific background variables. In practice of testing, some individual test items might have different psychometric quality for certain sub-populations, like boys/girls, urban/rural or private school/public school. But there are two good reasons not to include such differences in the research model:

1: in this research the data used and the correlations that can be made with sufficient statistical significance, must be aggregated to the levels of subtests, complete test and subjects in grade 7.

2: research into differences in the performance of individual items, based on some background variables – called Differential Item Functioning – are already part of the yearly analysis of selection test results.

The method used, LISREL (LInear Structural RELations), is an example of a structural equation model (Igolkina & Meshcheryakov, 2020). Structural equation modeling (SEM) includes a diverse set of mathematical models, computer algorithms, and statistical methods that fit networks of constructs to data. Structural equation models are often used to assess unobservable 'latent' constructs. They often invoke a measurement model that defines latent variables using one or more observed variables, and a structural model that imputes relationships between latent variables. The links between constructs of a structural equation model may be estimated with independent regression equations or through more involved approaches such as those employed in LISREL (Lei & Wu, 2007; Davvetas et al., 2020).

Data used in the research. Based on the data as provided for the cohort grade 7 in 2019-20, the research team developed a 'grand research database' in the statistical program SPSS (Statistical Product and Service Solutions; IBM). From this database the following data were defined:

Table 3: *Latent variables*

1: 18 latent variables on individual student performance in NIS Selection and grade 7

SEL19	M19MATH	M20MATH	M19ENG	M19K2/R2		
ENG	BIO	WHIST	GEO	CS	ART	
KHIST	KLIT/RLIT*	KZ/RU*	MATH	K2/R2**	PHYS	CHEM

* each student will have either Kazakh or Russian as first language and also for the same language the subject of Literature

** language and literature for second language

The overall performance of each individual student in the 2019 selection test was taken as one latent variable. This 'SEL19' variable is an important parameter of the capacities of the student indicating his or her readiness for a successful career in NIS.

Two main sources were used to obtain the data on academic achievement of the grade 7 students:

A:Monitoring assessment administrations in September and January:

- Mathematics first measurement at the start of the year: M19MATH
- Mathematics second measurement halfway the year: M20MATH
- Second/foreign Languages first measurement at the start of the year: M19K2/R2 and M19ENG

Unfortunately, NIS was not able, due to the COVID19 restrictions, to administer the second monitoring moment for languages in April 2020. These data are missing in this research.

B: the other 13 latent variables are the cumulated term and period scores in each subject in the grade 7 calendar plan, for which NIS provided assessment data.

ENG = English / BIO = biology / WHIST = World History / GEO = Geography / CS = Computer Sciences / ART = Arts / KHIST = Kazakh history / KZ = Kazakh language / MATH

= Mathematics / RULIT = Russian Literature / PHYS = Physics / CHEM = Chemistry

2: in total 73 variables could be defined as parameters for students' achievements in all the available data. These 73 variables are the subcategories of one of the 18 latent variables. For example, the latent variable SEL19 can be subdivided in 5 subscores, for each of the five subtests:

Table 4: *Five subtests*

ko_math	Subtest Mathematics; max. 400
ko_qr	Subtest QR; max. 300
ko_kz	Subtest Kazakh language; max. 200
ko_ru	Subtest Russian language; max. 200
ko_en	Subtest English language; max. 200

For monitoring scores, the five domains of mathematics or the four skills in a language are defined as being the subcategories for these specific latent variables.

For the 13 subjects the subcategories were defined by the four periods in the academic year for which these were collected. An example for the four terms in Geography: GEO_I, GEO_II, GEO_III and GEO_IV.

Results. To get a first indication of possible correlations between sub scores, simple linear correlations between the 18 latent variables were calculated. All correlation results were separately analysed for the Kazakh stream and for the Russian stream.

Kazakh stream. Correlation coefficients vary from very high - >0.80 - to quite low - <20 or even negative. So, there might certainly be a relation between classroom performance for specific subjects and overall selection test performance. In other words, the results are supportive for a possible relation between selection test and some of the (subject) school results. In the table 3 we summarise the strongest and the weakest correlations.

Table 5: *Combinations of selection test result and school subjects with relatively high and relatively low correlations; Kazakh stream*

High correlation with SEL19 (more than 0.5)	Low correlation with SEL19 (less than 0.2)
Mathematics monitoring 7.1	Arts
Mathematics terms I - IV	Kazakh literature
Mathematics monitoring 7.2	Kazakh first language
	Kazakh history
	Geography
	English (0.18)

These results show a strong positive correlation between the overall selection test result and all Mathematics assessments in grade 7. More surprisingly might be the low correlation for the language subjects Kazakh and English and other subjects that are taught in Kazakh language. And also a correlation of just 0.19 for English language performance seems to be lower than one might expect.

Low correlation scores as such are not necessarily a 'bad sign'; in the case of some languages relative low correlations between SEL19 and the grade 7 results might be explained by:

1: the difference between what is assessed in SEL19 and what is taught, learned and assessed in grade 7 curriculum and practice. Selection test is only assessing reading comprehension, whereas in grade 7 the languages also teach and assess writing, listening and speaking;

2: the positive effect of learning progress that students make once they enter NIS grade 7.

These low correlation scores are seen with the Kazakh language subjects and less or on a less low level with the Russian language subjects.

This difference between correlation scores for both languages demands for further research, quantitative and qualitative (interviews, focus groups).

Russian stream. In Table 4, we provide a summary of the strongest and the weakest correlations for Russian stream.

Table 6. *Combinations of selection test result and school subjects with relatively high and relatively low correlations; Russian stream*

High correlation with SEL19 (more than 0.5)	Low correlation with SEL19 (less than 0.2)
Mathematics monitoring 7.1	Arts
Mathematics terms I - IV	Kazakh literature
Mathematics monitoring 7.2	Kazakh history
Physics terms I - IV	English (0.27)

In comparison to the Kazakh stream, the Russian stream shows a strong positive correlation between the overall selection test result and all mathematics assessments in grade 7. Also, not very surprisingly might be the very low correlation for the subjects that are closely related to Kazakh language. And even a correlation of just 0.27 for English language performance (0.19 it is for the Kazakh stream) seems to be lower than one might expect for a subject that partly decides on NIS entrance.

Discussions. The 2020 validation research confirmed that the NIS selection test, despite modifications in 2019, continues to show strong predictive validity, particularly in mathematics and quantitative reasoning. However, the moderate predictive accuracy in languages reflects a broader trend in education where language subtests often show weaker predictive power compared to mathematics (Dale & Sparks, 2024).

This trend aligns with findings by Jamali et al. (2023), who note a global focus on improving educational quality and identifying gifted students. A similar approach is observed in Denmark's educational policies, which aim to support talented students across all educational levels (Rasmussen & Lingard, 2018).

Despite these global efforts, many empirical studies overlook how various abilities in gifted students interact and affect their development, particularly the link between language proficiency and STEM learning (Boer & Rijnsoever, 2022). However, existing literature shows positive correlations between language proficiency and STEM learning, as demonstrated by MacGregor (1999), Vukovic (2013), Prediger (2013), and Hernandez (2013). Crossley (2018) further suggested that high mathematics scores could positively influence language test results.

The introduction of the natural sciences subtest in 2022 was also a logical step, stemming from the 2020 validation research findings. This new subtest, which includes 20 items covering Physics, Chemistry, Biology, and Geography, expanded the test's scope and reflects ongoing efforts to improve its predictive validity.

Nevertheless, the language subtest (Day 2 booklet) remains an area of concern, and both the 2015 and 2020 validation studies recommend further research into its predictive power.

Conclusions. What do these correlation results tell us about the possible answers on the two main research questions:

1: Have the suitable students be selected?

2: Are the students selected in the right way?

The answer to the first question can only partly be based on the correlation findings from the cohort 2019-2020 performances in Selection 2019 with their assessment results in grade 7. Successfulness of those learners who received a grant for studying at NIS can be measured by looking at their grade 7 results, but those who were not selected for grade 7 cannot be included in comparable research. We do not have data on their grade 7 results, as they did not participate. Some of them, especially those who were close to selection in the 2019 ranking, might have

been quite successful learners in NIS grade 7. This is a category that statisticians call 'the false negatives': candidates who would have had the ability to study successful in grade 7 at NIS but were not selected because their overall score on SEL19 was not high enough to be within the range of available vacant places.

The answer to the second question can partly be answered by looking at the outcomes of the correlation research. The highest correlations between score on selection test 2019 and marks achieved in grade 7 are realized for the subtests Mathematics and QR. These are also the two subtests for which the Selection scoring uses a minimum passing score. This pass score gives the subjects Mathematics and QR a relative greater power in the overall decision for student selection. These two subtests also contribute to more than 50% of the total overall score for selection test (400+300 out of 1300).

In general, we might state that, as far as correlation indices can be informative for answering this question, the selection test has proven to be effective in selecting the candidates for entering NIS grade 7 in a reliable and valid way.

References:

Boer, T., & Van Rijnsoever, F. (2022). In search of valid non-cognitive student selection criteria. Assessment & Evaluation in Higher Education, 47(5), 783-800. https://www.tandfonline.com/doi/epdf/10.1080/02602938.2021.1958142?needAccess=true

Center for Pedagogical Measurements (2023). Competitive selection. https://cpi-nis.kz/en/

Chomphuphra, P., Chaipidech, P., & Yuenyong, C. (2019). Trends and research issues of STEM education: A review of academic publications from 2007 to 2017. Journal of Physics: Conference Series (Vol. 1340, No. 1, p. 012069). IOP Publishing. https://iopscience.iop.org/article/10.1088/1742-6596/1340/1/012069/pdf

Crossley, S., Ocumpaugh, J., Labrum, M., Bradfield, F., Dascalu, M., & Baker, R. S. (2018). Modeling Math Identity and Math Success through Sentiment Analysis and Linguistic Features. International Educational Data Mining Society. https://files.eric.ed.gov/fulltext/ED593117.pdf

Dale, P. S., & Sparks, R. L. (2024). Mapping the predictive role of MLAT subtests for L2 achievement through regression commonality analysis. *Studies in Second Language Acquisition*, *46*(2), 453-477.

Davvetas, V., Diamantopoulos, A., Zaefarian, G., & Sichtmann, C. (2020). Ten basic questions about structural equations modeling you should know the answers to–But perhaps you don't. Industrial Marketing Management, 90, 252-263. https://www.sciencedirect.com/science/article/abs/pii/S0019850120302935?via%3Dihub

Hernandez, F. (2013). The Relationship Between Reading and Mathematics Achievement of Middle School Students as Measured by the Texas Assessment of Knowledge and Skills (Doctoral dissertation). https://uh-ir.tdl.org/items/45d2a1e2-6ba4-4dc5-81a4-99675dfd8d64

Igolkina, A. A., & Meshcheryakov, G. (2020). semopy: A python package for structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 27(6), 952-963. https://www.tandfonline.com/doi/full/10.1080/10705511.2019.1704289?scroll=top&needAccess=true

Jamali, S. M., Ale Ebrahim, N., & Jamali, F. (2023). The role of STEM Education in improving the quality of education: A bibliometric study. International Journal of Technology and Design Education, 33(3), 819-840. https://link.springer.com/article/10.1007/s10798-022-09762-1

Lei, P. W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. Educational Measurement: issues and practice, 26(3), 33-43. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1745-3992.2007.00099.x

MacGregor, M., & Price, E. (1999). An exploration of aspects of language proficiency and algebra learning. Journal for Research in Mathematics Education, 30(4), 449-467. https://www.jstor.org/stable/749709?origin=crossref

Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions—design and effects of a language-and mathematics-integrated intervention. Mathematics Education Research Journal, 25(3), 435-456. https://link.springer.com/article/10.1007/s13394-013-0079-2

Rasmussen, A., & Lingard, B. (2018). Excellence in education policies: Catering to the needs of gifted and talented or those of self-interest? European Educational Research Journal, 17(6), 877-897. https://journals.sagepub.com/doi/epub/10.1177/1474904118771466

Vukovic, R. K., & Lesaux, N. K. (2013). The relationship between linguistic skills and arithmetic knowledge. Learning and Individual Differences, 23, 87-91. https://www.sciencedirect.com/science/article/abs/pii/S1041608012001471

Назарбаев Интеллектуальные школы (2023). Критериальное оценивание. https://www.nis.edu.kz/en/kriterialnoe

Назарбаев Интеллектуальные школы (2023). Правила конкурсного отбора. https://www.nis.edu.kz/storage/app/media/konkursny%20otbor/pravila%20konkursnogo%20otbora/ministra-prosveshcheniya-rk-255.pdf

References:

Boer, T., & Van Rijnsoever, F. (2022). In search of valid non-cognitive student selection criteria. Assessment & Evaluation in Higher Education, 47(5), 783-800. https://www.tandfonline.com/doi/epdf/10.1080/02602938.2021.1958142?needAccess=true

Center for Pedagogical Measurements (2023). Competitive selection. https://cpi-nis.kz/en/

Chomphuphra, P., Chaipidech, P., & Yuenyong, C. (2019). Trends and research issues of STEM education: A review of academic publications from 2007 to 2017. Journal of Physics: Conference Series (Vol. 1340, No. 1, p. 012069). IOP Publishing. https://iopscience.iop.org/article/10.1088/1742-6596/1340/1/012069/pdf

Crossley, S., Ocumpaugh, J., Labrum, M., Bradfield, F., Dascalu, M., & Baker, R. S. (2018). Modeling Math Identity and Math Success through Sentiment Analysis and Linguistic Features. International Educational Data Mining Society. https://files.eric.ed.gov/fulltext/ED593117.pdf

Dale, P. S., & Sparks, R. L. (2024). Mapping the predictive role of MLAT subtests for L2 achievement through regression commonality analysis. *Studies in Second Language Acquisition*, *46*(2), 453-477.

Davvetas, V., Diamantopoulos, A., Zaefarian, G., & Sichtmann, C. (2020). Ten basic questions about structural equations modeling you should know the answers to–But perhaps you don't. Industrial Marketing Management, 90, 252-263. https://www.sciencedirect.com/science/article/abs/pii/S0019850120302935?via%3Dihub

Hernandez, F. (2013). The Relationship Between Reading and Mathematics Achievement of Middle School Students as Measured by the Texas Assessment of Knowledge and Skills (Doctoral dissertation). https://uh-ir.tdl.org/items/45d2a1e2-6ba4-4dc5-81a4-99675dfd8d64

Igolkina, A. A., & Meshcheryakov, G. (2020). semopy: A python package for structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 27(6), 952-963. https://www.tandfonline.com/doi/full/10.1080/10705511.2019.1704289?scroll=top&needAccess=true

Jamali, S. M., Ale Ebrahim, N., & Jamali, F. (2023). The role of STEM Education in improving the quality of education: A bibliometric study. International Journal of Technology and Design Education, 33(3), 819-840. https://link.springer.com/article/10.1007/s10798-022-09762-1

Lei, P. W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. Educational Measurement: issues and practice, 26(3), 33-43. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1745-3992.2007.00099.x

MacGregor, M., & Price, E. (1999). An exploration of aspects of language proficiency and algebra learning. Journal for Research in Mathematics Education, 30(4), 449-467. https://www.jstor.org/stable/749709?origin=crossref

Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions—design and effects of a language-and mathematics-integrated intervention. Mathematics Education Research Journal, 25(3), 435-456. https://link.springer.com/article/10.1007/s13394-013-0079-2

Rasmussen, A., & Lingard, B. (2018). Excellence in education policies: Catering to the needs of gifted and talented or those of self-interest? European Educational Research Journal, 17(6), 877-897. https://journals.sagepub.com/doi/epub/10.1177/1474904118771466

Vukovic, R. K., & Lesaux, N. K. (2013). The relationship between linguistic skills and arithmetic knowledge. Learning and Individual Differences, 23, 87-91. https://www.sciencedirect.com/science/article/abs/pii/S1041608012001471

Nazarbayev Intellectual schools (2023). Kriterial'noe otsenivanie [Criteria-based assessment]. https://www.nis.edu.kz/en/kriterialnoe [in Russian]

Nazarbayev Intellectual schools (2023). Pravila konkursnogo otbora [Rules of competitive selection]. https://www.nis.edu.kz/storage/app/media/konkursny%20otbor/pravila%20konkursnogo%20otbora/ministra-prosveshcheniya-rk-255.pdf [in Russian]