

THE DEVELOPMENT OF TEACHERS' ASSESSMENT OF PUPILS DURING PRIMARY SCHOOL

Irena Smetáčková¹, Petr Pavlík²

¹ *Pedagogická fakulta, Univerzita Karlova*

² *Fakulta humanitních studií, Univerzita Karlova*

DOI: 10.21062/edp.2023.001

Abstract

How teachers perceive and assess students has a key impact on the success of student learning. Teachers' assessments of students can take many forms - grades awarded, explicit perceptions of school success, or more detailed perceptions of school characteristics that may promote the chance of successful mastery of demands. If teacher assessment is consistent across these forms, it can provide better feedback, motivation and accountability for their own learning. However, consistency is also important across time. If assessment varies significantly over time, it may be less comprehensible to pupils. In addition, it may (or may not) be indicative of unclear criteria used by teachers in assessment. The study presented here focused on the consistency of teachers' assessment of specific pupils across forms of assessment and across time. At two-year intervals, teachers were asked to comment explicitly on the school performance of individual children (whether they were successful or not), to complete an assessment scale containing 32 school characteristics in six domains, and finally to provide data on the progress received by pupils and estimated for them in subsequent years. The research sample consisted of 7 classes with a total of 171 pupils who were repeatedly assessed by their teachers. Analyses showed that there was relative agreement in the forms of teacher assessment across the measures. Teachers assessed pupils in a rather positive way. However, assessments of 32 characteristics showed more variability and were thus more informative than grades and school performance assessments. Across the measures, the assessments showed significant deterioration. For Year 4 pupils, teachers were more critical in their assessments, and again, the comprehensive assessment of school characteristics was more varied than the assessment via grades and via explicit assessment of school achievement.

Klíčová slova: school success, school failure, primary school, assessment, evaluation, grades

INTRODUCTION

School failure is usually defined as poor academic performance that does not meet an expected competence and falls below the desired standards. This definition puts teachers in the center because the expected standards are linked to the requirements given or at least represented by teachers. Many studies, including the one carried out by our team (Smetackova, Stara and Chytry, 2023), showed that individual teachers vary in their subjective descriptions of what it means to be successful or failing in school. It corresponds with the study by Rubie-Davies (2010) who found that the average levels of expectations for students in the classroom are different. The reason is deeply rooted in teachers' subjective theories that represent a set of beliefs allowing teachers to understand their educational environment and themselves (Groeben & Scheele, 2000). However, Ready and Wright (2011) found that roughly half of disparities between teacher's perception of literacy skills and students' background are explained by actual between-group differences and the other half results more from classroom characteristics than from teachers' professional or personal backgrounds.

Teachers' views of students have a strong impact on students' learning and education outcomes. Teachers create the structure of learning situations, motivate activity, provide feedback, and make decisions about the future direction of students (Rubie-Davies, 2008). In doing so, they influence how student learn, how responsible they are for their own learning, and what their educational aspirations are (Südkamp, Kaiser & Möller, 2012). They also influence students' self-efficacy. If teachers have accurate and fitting understanding of school characteristics of individual pupils they can provide adequate support for their learning.

Vice-versa, inaccurate understanding leads to practices that do not correspond to the students' educational potential and needs. It can hinder students' academic performance (Rubie-Davies, 2008; Südkamp, Kaiser & Möller, 2012). The impact of teachers' expectations has been confirmed by many studies. For example, Hattie (2012) found a moderate average effect of teacher expectations on student performance. Van den Boer et al. (2010) confirmed that negatively biased expectations correlated with low student achievements, while positively biased expectations correlated with high student achievements.

Jussim and Harber (2005) claim that teachers' expectations are rather accurate because they are built on perceptions of previous achievements. However, Rubie-Davies (2008), Südkamp et al. (2012) and others disagree. They suggest that expectations can serve as the self-fulfilling prophecy which might decrease the efficiency of learning process of individual students (Rist, 2017). However, if teacher's views are accurate, the risk of self-fulfilling prophecy is low. The intervention study by Gatlin-Nash et al. (2021) shows that accuracy of teacher's perceptions of students' academic competences might be improved.

Teachers' perceptions of students are transformed into pedagogical assessment. This can take different forms. There are explicit and highly structured forms such as grades. This is summative assessment. There are also implicit and low-structured forms such as teachers' ideas about the school characteristics of individual pupils. In this case, it is usually formative assessment.

We still need more research to better understand the complexity of teachers' view on students, especially at the beginning of schooling. Two aspects need to be examined in more details: 1) correspondence among different teachers' assessments of individual pupils, and 2) development of teachers' views across time (Jussim & Harber, 2005; Timmermans, de Boer & van der Werf, 2012).

PRESENT STUDY

The present study is a part of the research Teachers' Understanding of the Causes of School Failure at the Start of Compulsory Education¹. The aim of the research was to identify pedagogical practices used by teachers towards children at risk of school failure in primary school classrooms. Besides other data, teachers' perceptions of individual pupil's abilities and performance were collected through the rating scale. In two years, the same rating scale was used to compare whether and how teachers' perceptions have evolved. The present study answers the following research questions: *How do teachers' assessments of pupils evolve two years apart? Which areas of pupil dispositions, performances and conditions are evaluated differently over time?*

Sample

The sample of the present study consisted of 171 pupils, including 76 boys (44%) and 95 girls (56%). Pupils belonged to 7 primary school classes in which data were collected in Year 2 and Year 4. These were regular classrooms of 22 to 28 pupils located in different parts of the Czech Republic, in municipalities of different sizes (small villages, smaller and medium-sized towns).

The classes were part of a set of 29 classes that participated in the main research. The classes were approached through two channels - about 1/3 of the teachers had previously collaborated with the research team, another 1/3 were recommended by the school management who had previously collaborated with the team, and the remaining 1/3 expressed interest in collaboration based on randomly sent out outreach letters. Parents' and children's consent for long-term participation in the research as well as the teachers' excellent teaching reputation were conditions for inclusion of specific classes in the research. The latter was inferred from the fact that school administrators and randomly interviewed parents of children in the upper grades agreed that they would recommend the teacher and would like him/her to teach their own children. This condition was imposed because the aim of the research was to find out what practices against school failure are used by 'good' teachers.

Methods

¹ The research is carried out by the Faculty of Education of Charles University with the financial support of the Ministry of Education (project OPVV, project number CZ.02.3.68/0.0/0.0/19_076/0016390).

² Twenty-nine classrooms participated in the main study, with data collected in Year 2. In only seven classes did the same teachers remain until Year 4, and therefore the development of their perceptions of pupils could be ascertained and compared.

In the present study, three kinds of data were analysed: a) expert assessment of pupil's characteristics through the teacher's questionnaire "Pupil Characteristics" in Years 2 and 4, b) explicit assessment of pupil's achievement, and c) grades.

The "Pupil Characteristics" questionnaire included 32 short statements that described the child's school performance and the key factors involved. The statements related to pre-selected criteria that were identified from the literature and teacher interviews. The interviews also helped to identify the specific wording of the statements. Before being used in the research, the questionnaire was piloted through qualitative interviews with 5 teachers outside the research population and quantitative analysis in two classrooms.

The statements were divided into the following six subscales: 1) cognitive abilities, 2) other abilities, 3) relationships, 4) school performance, 5) emotions, 6) family background. The questionnaire is presented in Table 1.

The reliability of the questionnaire measured by Cronbach's Alpha was sufficient. It was .935 for the Year 2 and .920 for the Year 4. The reliability of individual subscales was lower, but still sufficient. It was .85 and .88 for the Cognitive abilities subscale; .75 and .57 for the Other abilities subscale; .76 and .89 for the Relationships subscale; .83 and .83 for the School performance subscale; .61 and .45 for the Emotions subscale; .88 and .77 for the Family background subscale. The only problematic subscale is the Emotions with too low measures.

For each statement, teachers evaluated whether it applies (1 point), applies partially (2 points), or does not apply (3 points) to the pupil. The direction of the rating scale was chosen to match the classification norm - i.e. a child with positive characteristics received a "grade" of 1. For each child, the teacher's rating was used to calculate the total score. The fewer points (i.e. lower scores), the more positive school characteristics the teachers think the child has (i.e. the more successful he/she is in school).

Teachers also assessed individual children through the statement "Is he or she successful in school?", which followed the questionnaire "Characteristics of the pupil". The assessment used the same responses, i.e., the statement applies (1 point), applies partially (2 points), or does not apply (3 points).

For individual pupils, we also obtained data on their grades in mathematics and Czech language at the end of Year 2 and in the middle of Year 3. In the Czech educational system, the grades from 1 (best) to 5 (worst) are used. From this, the average value of the actual grades obtained was calculated. We also collected the teacher's estimate of the pass mark at the end of Year 3 and Year 5 in mathematics and Czech language. From these, the average value of the estimated grades at the end of the primary school were calculated.

PRESENT STUDY

School grades

The grades which the children received in Years 2 and 3 in mathematics and Czech language were very good. The average of the real grades awarded was 1.157 (SD=.334). More than 80% of the pupils always obtained an excellent mark. In Year 3, there was a slight deterioration in grades. While in Year 2 and Year 3 only 9% and 13% of pupils respectively received grades other than A, the percentage increased to 17% in Year 3. The trend of worsening grades may be the result of a higher level of difficulty in the curriculum or more stringent assessment criteria.

Comparison with the estimated grades in the following years confirmed the trend of deterioration. The mean of the estimated grades at the end of Year 3 and Year 5 was 1.359 (SD=.439). The difference is statistically significant ($p<.001$). Teachers estimate that at the end of Year 5 only 45% of pupils will have an excellent grade in Czech language and 48% in mathematics. They expect a grade 3 for 11-12% of pupils.

School success (explicitly)

Teachers' evaluation of pupils was also rather positive. Explicit comments on school success (i.e., assessing the validity of the statement "He/she is school successful") revealed that very few children were not considered successful by teachers. In Year 2, only 3 children were identified as not successful and 29 as partially successful. In Year 4, 3 children were also identified as not successful and 34 as partially successful. The increase in perceived school failure over time was negligible (compared to grades with significant deterioration).

In total, there were 118 children in the cohort who were consistently rated by their teachers as academically successful. Twelve children were considered partially successful in both tests and 37 children in one test. Two children were identified as not successful in both tests and another two children in one test. For the sake of

clarity, we divided the sample into children perceived as successful (n=118) and children perceived as partially or completely unsuccessful (n=53).

Assessment of school characteristics

Teachers' evaluations of the children in the Pupil Characteristics Questionnaire on the scale 1-3 were also quite positive. The overall mean score in Year 2 was 1.252 and in Year 4 was 1.319. The difference between the two measures was statistically significant ($p < .001$) and indicated a decrease in teacher evaluation. Older children were more likely to be rated by the teachers as less proficient in each aspect (i.e. positive statements were more likely not to apply).

All 32 assessed statements are presented in Table 1. It also contains the average rating of individual statement in the first and the second measurements, as well as a statistical comparison of the significance of the difference (Paired Sample t-Test). For a half of the statements, the difference was statistically significant, of which 14 showed a worsening of the rating and only two showed an improvement (helping others and self-care).

Table 1: Expert assessment in the questionnaire "Pupil characteristics"

Sub-scale	Statements	2nd grade		4th grade		p
		Mean	SD	Mean	SD	
Cognitive abilities	He/she can concentrate for 10-15 minutes on one activity.	1.26	.577	1.45	.704	< .001
	He/she has a good memory, remembers easily.	1.36	.610	1.50	.723	.006
	He/she is a logical thinker.	1.39	.577	1.57	.703	< .001
	He/she is inquisitive, likes to learn new things.	1.31	.578	1.41	.603	.035
	He/she immediately understands the task and the teacher's instructions.	1.52	.689	1.54	.697	.702
Other abilities	He/she is diligent and hardworking.	1.36	.602	1.58	.718	< .001
	He/she is creative, has its own ideas.	1.33	.614	1.55	.721	< .001
	He/she is manually dexterous.	1.31	.587	1.26	.513	.199
	He/she is physically fit and good at sports.	1.31	.587	1.36	.590	.218
	He/she can do all the necessary self-care.	1.08	.315	1.02	.132	.007
Relationships	He/she is integrated into the class team and has good relations with other children.	1.09	.330	1.16	.425	.048
	He/she can work together in pairs or small groups.	1.15	.386	1.15	.443	1.00
	He/she is helpful to others.	1.22	.494	1.13	.369	.018
	He/she communicates well with me (the teacher).	1.08	.296	1.13	.385	.088
	He/she communicates well with other adults in the school, e.g. the teaching assistant.	1.08	.290	1.10	.319	.395
	He/she is not afraid to ask if something is not clear to him/her or does not know something.	1.16	.366	1.23	.498	.058
School performance	In the first grade, he/she is adapted well to school.	1.12	.375	1.21	.485	.007
	He/she is good at counting (math).	1.28	.476	1.41	.601	< .001
	He/she is good at writing.	1.29	.571	1.31	.556	.727
	He/she is good at reading.	1.29	.514	1.30	.563	.774
	He/she has knowledge and insight beyond the school curriculum.	1.65	.763	1.73	.797	.637

	He/she can work with a mistake, learn from it.	1.46	.644	1.49	,698	.580
	He/she is active and involved in his lessons.	1.36	.551	1.53	,617	.002
	He/she works quickly, with tasks soon completed.	1.71	.802	1.86	,746	.012
Emotions	He/she likes going to school.	1.12	.327	1.38	,586	< .001
	He/she responds to success with visible joy.	1.15	.456	1.26	,524	.005
	He/she has appropriate reactions to failure.	1.11	.392	1.06	,310	.162
	He/she is emotionally mature for her age.	1.11	.377	1.09	,340	.603
Family background	Parents communicate well with the school.	1.06	.246	1.06	,310	1.00
	Parents help him/her adequately with the schoolwork.	1.08	.328	1.16	,452	.023
	He/she has a working family background.	1.08	.329	1.12	,419	.181
	Parents have realistic ideas of what the child can achieve.	1.09	.405	1.12	,360	.619

The evaluated statements were divided into six areas. Due to the space limitations, it is not possible to describe the process of creating the areas and therefore we refer to other texts. For each domain, the average score was calculated and compared for Years 2 and 4. The results, including the statistical significance, are presented in Table 2.

Table 2: Expert assessment in the questionnaire "Pupil characteristics" by area

	2nd grade		4th grade		p
	Mean	SD	Mean	SD	
Cognitive abilities	1.366	.480	1.498	.575	< .001
Other abilities	1.279	.392	1.353	.349	< .001
School performance	1.396	.401	1.478	.433	.006
Relationships	1.131	.247	1.148	.327	.460
Emotions	1.121	.264	1.197	.279	< .001
Family background	1.084	.274	1.120	.334	.047
TOTAL	1.252	.287	1.319	.306	< .001

Both the total score and the sub-scores in five areas showed significant differences. Four areas showed deterioration, one (emotional maturity) showed improvement. However, the subscale Emotions has too low reliability which needs to be considered in the interpretation. In the areas of cognitive ability, other skills necessary for successful school adjustment and school performance, the deterioration in teacher ratings was statistically significant for older children. The deterioration in ratings of family background was on the edge of statistical significance.

School characteristics, school grades and explicit school success/failure

We divided the sample into two groups based on the explicit assessment of school success - children perceived as successful (n=118) and children perceived as partially or completely unsuccessful (n=53). We then compared these across teachers' ratings of sub-characteristics of individual pupils to see what areas most saturated teachers' explicit evaluations of pupils as unsuccessful and whether this pattern changed over time.

Pupils identified as school achievers scored significantly higher on both measures, both across the whole questionnaire and in all the areas of the questionnaire. The overall mean teacher rating in Year 2 for pupils with perceived school success was 1.127 (SD=.165), while for pupils who were partially or completely unsuccessful it was 1.531 (SD=.308). In Year 4, the results were similar at 1.184 (SD=.179) and 1.622 (SD=.314), respectively.

These results suggest that teachers' ratings of overall school success are based on perceived failure in certain sub-areas. Thus, the teacher's perspective is internally consistent.

Another piece of information about whether teachers consider students to be successful in school is their grades. We were interested in the extent to which achievement is related to detailed assessment of pupil characteristics. To this end, we conducted a correlation analysis (Pearson Correlation). It revealed a significant correlation between the grades and the total score on the Pupil Characteristics Questionnaire for both measurements. For the first measurement, the correlation was .460, for the second measurement .564. The strongest correlations with grades were the school performance assessment ($r=.518$ and $.608$), the cognitive ability assessment ($r=.423$ and $.521$), and the family background assessment ($r=.522$ and $.403$). The correlations with the other domains were also statistically significant, but only around .2.

Comparison of classes

Seven classrooms participated in the study, where pupils were assessed always by one teacher. It is, therefore, also interesting to look at the consistency of assessment at the level of individual classes. The analysis showed strong variability among classes. In two classes, the differences in scores across the first and second measures were not statistically significant. However, in the other classes, they were significant, with one case showing an improvement in scores and four showing a deterioration in scores. Although the largest changes were observed most often with respect to cognitive ability and school performance (which is consistent with the overall results), this was not a general pattern. For example, in one class, the strongest changes were in the assessment of family background, while in another class, no change was observed in this area. Similarly, school performance, which was one of the strongest areas of change in most classes, showed smaller change than the assessment of cognitive ability in two classes. These differences across classes suggest that individual teachers apply different criteria in their assessments and factor in the increase in school demands differently.

DISCUSSION

The present study focused on the consistency of teachers' perception of specific pupils across time and across forms of assessment. Explicit representations of school achievement, grades and comprehensive assessment of school characteristics are indicative of teachers' perspectives of students from different angles. We hypothesized that they should coincide. This was only partially confirmed. The assessments followed the same direction - they were rather positive in all three cases. However, the degree of homogeneity in each type of assessment was different.

The assessment of 32 characteristics showed the greatest variability, while the other two assessments were less sensitive to diversity among pupils. This may have been a consequence of lesser complexity of the assessments - both the achievement and the explicit assessment of school performance are too general tools and tend to use fewer dimensions. It could also be because these are summative assessments with respect to which the teachers feel a risk of certain fatality and therefore tend to be less strict. This was confirmed in this research also by the analysis of interviews with teachers which is presented in another article (Smetáčková, Stará and Chytrý, 2023).

The finding that there is only partial coincidence between the three forms of teachers' assessment suggests that teachers' assessments are not based on sufficiently clarified criteria that would be applied across different situations. Thus, we dispute conclusions of Jussim and Harber (2005) who argue that teachers' assessments are accurate. If they were, coincidence of individual forms of assessment should be stronger.

The strongest relations were found between grades and cognitive abilities, other abilities and family backgrounds. The high impact of cognitive abilities on school achievements were confirmed by many studies (e.g. Spinath et al., 2006). However, in our study, the correlation with other abilities and family background was even stronger.

With regard to the development of the assessment, the second measure showed a greater variability as well as a greater criticality of teachers' perspectives. The fact that teachers assess older children as less successful may seem paradoxical. These are still the same children whose innate abilities remain the same. It might be expected that pupils develop in their school performance or relationships, but not in cognitive abilities. Yet the greatest changes were recorded precisely in those items that are described as given and unchanging in the usual discourse, and they were in the direction of deterioration. This is surprising given that children at school gradually develop their knowledge and skills. Teachers could/should therefore assess them likewise (in the case of innate resources) or even better (in the case of learned resources). The contradictory result indicates that teachers

tighten gradually their criteria and abandon the primarily positive-motivational function of feedback (assessment becomes more of a source of complex feedback with age).

A significant deterioration in the assessment could also result from insufficient accuracy of expert assessments after the first year. If teachers had been able to make more accurate pedagogical diagnoses, the difference in assessment over time would not have been so pronounced (i.e. they would have been able to identify more pupils with weaknesses in the aspects under study already in the first measurement). It could be argued that the increasing difficulty of school tasks gradually reveals limits of individual children, which may have not been apparent at the beginning of schooling. However, this would not be the case if teachers made greater use of a dynamic approach to diagnosis and graded demand factor of school tasks. Even though our sample consisted of teachers with very good pedagogical competences and excellent expert reputation, the incidence of such approaches was very low.

An alternative view can be added to the dominant explanation of the results presented above. It would refer to a self-fulfilling prophecy (Rist, 2017). Teachers were aware that some Year 2 pupils had weaknesses with respect to certain school characteristics, but this did not yet translate to lower summative scores because of motivational reasons mentioned above. However, due to the self-fulfilling prophecy, this negative perception could intensify, and in Year 4, it was already reflected in summative assessments. These pupils were perceived as school failures and it was expected that they would perform worse.

Such an explanation would be in line with other studies which confirm the influence of teachers' beliefs on students' future school development (e.g., Südkamp et al., 2012; van de Boer et al., 2010). Our data do not allow us to confirm or dismiss the self-fulfilling-prophecy hypothesis. We recommend focusing attention on this in future research. However, the high inter-teacher variability found in our sample would be theoretically more in line with the confirmation of the self-fulfilling prophecy. At the same time, differences in perceptions of pupils among teachers are also found in other studies, e.g. Rubbie-Davies (2008) or Ready and Wright (2011).

The study has several limitations. The first concerns the instrument itself. Although the scale has undergone standard development and shows good psychometric characteristics, some subscales (specifically Emotions) had rather low reliability. Thus, caution should be exercised in interpreting the results of individual subscales. It is hoped that the scale will be used in other research that will help to improve its quality. The second limitation is the small sample size (N=171). The findings should be tested on a larger number of pupils. However, we recall that the implementation of longitudinal research is very challenging and therefore we consider our sample of respondents to be successful. A third limitation is the data sources available to us in the analysis. To explain the quantitative difference found in the first and second assessments, it would have been useful to also conduct interviews with the teachers. A mixed design study would have been beneficial. The last limit is the sample of teachers. All of them were perceived by their school principals, colleagues, and parents of pupils as educational experts. The findings about high consistency of teachers' assessment criteria should be examined in different group including teachers with shorter pedagogical practice and lower competencies.

CONCLUSION

The teacher's view of the pupils is crucial in the learning process. The teacher sets up learning situations, motivates, gives feedback, and carries out pedagogical assessment. All this is influenced by the teacher's perception of the pupil. Part of professional development is for teachers to learn to form their perceptions of pupils because of adequate resources and thus avoid errors in social perception and cognition that could lead to bias. Indeed, if a teacher has an incorrect perception of a pupil, the learning process and personality development of the pupil can be severely impaired. Teachers often believe that if they wish to achieve realistic perceptions of their pupils, this alone is enough and therefore that this wish alone makes their perceptions sufficiently accurate. However, this may be a mistaken belief. Research that shows that there are errors in teachers' perceptions of pupils, which teachers make them and for what reasons, is valuable. Not only for the sake of advancing scientific knowledge, but also as an argument for school practice that mere intention is not enough, and that the accuracy of teachers' views must be actively pursued.

Our study addressed the specific question of the development of teachers' perceptions over time. It produced relatively positive findings in relation to the criteria applied in pupils' perceptions and in formal assessment (grades). At the same time, however, the analysis showed a trend towards more rigorous assessment as pupils got older. This is of course understandable in view of the increasing demands of school and the retreat of purely positive motivation. However, this trend appears less understandable when assessing pupils' cognitive abilities, which should be stable. It is likely that the more critical view stems from inadequate assessment of cognitive

ability at the beginning of schooling. We suggest that this is the direction in which subsequent research and, where appropriate, practical recommendations for adjusting teacher preparation and professional development should be directed.

References

- [1] Gatlin-Nash, B., Hwang, J. K., Tani, N. E., Zargar, E., Wood, T. S., Yang, D., Powell, K. B. & Connor, C. M. (2021). Using assessment to improve the accuracy of teachers' perceptions of students' academic competence. *The Elementary school journal*, 121(4), 609-634.
- [2] Groeben, N., & Scheele, B. (2000, June). Dialog-Konsens-Methodik im Forschungsprogramm Subjektive Theorien. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 1(2), 2-00.
- [3] Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- [4] Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155.
- [5] Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48(2), 335-360.
- [6] Rist, R. C. (2017). On understanding the processes of schooling: The contributions of labeling theory. *Exploring Education*, 165-176.
- [7] Rubie-Davies, C. M. (2008). Teacher expectations. In T. Good (Ed.), *21st century education: A reference handbook* (pp. 1-254-1-265). Thousand Oaks: SAGE Publications. doi:[10.4135/9781412964012.n27](https://doi.org/10.4135/9781412964012.n27).
- [8] Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship? *British Journal of Educational Psychology*, 80(1), 121-135.
- [9] Smetackova, I., Stara, J., & Chytry, V. (2023). Teachers' perspectives on school failure/success in primary 1: the need for a new definition. *Studia Pedagogica*, (in press).
- [10] Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34(4), 363-374.
- [11] Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743-762. <https://doi.org/10.1037/a0027627>
- [12] Timmermans, A. C., Kuyper, H., & Van der Werf, M. P. C. (2015). Accurate, inaccurate or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *British Journal of Educational Psychology*, 85, 459-478. doi:[10.1111/bjep.12087](https://doi.org/10.1111/bjep.12087).
- [13] Van de Boer, H., Bosker, R. J., & van der Werf, M. P. C. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology*, 102, 168-179. doi:10.1037/a0017289