Contemporary Educational Psychology 59 (2019) 101787

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Exploring the Pygmalion effect: The role of teacher expectations, academic self-concept, and class context in students' math achievement

Grzegorz Szumski[a,1], Maciej Karwowski[b,1]

[a] Department of Education, University of Warsaw, Mokotowska 16/20 St., 00-561 Warsaw, Poland
[b] Institute of Psychology, University of Wroclaw, ul. Dawida 1, 50-527 Wroclaw, Poland

ARTICLE INFO

Keywords:
Teacher expectancy effects
Pygmalion
Inclusive classes
Math self-concept
Math achievement

ABSTRACT

Teacher expectancy effect (TEE) also known as the Pygmalion effect is a classic, yet still controversial phenomenon within educational psychology. In this paper, we examine TEE in a longitudinal study on a large sample (N = 1488) of Polish middle-school students and their teachers. Consistent with TEE, teachers' higher expectations were positively related to students' math achievement three semesters later, even after controlling for initial achievement. Students' academic self-concept in math partially mediated the observed effect: teachers' higher expectations translated into higher students' academic self-concept, which consequently predicted their higher math achievement. Importantly, TEE was observed not only on the level of individual students, but also on class level. Higher expectations of the entire class improved individual achievements of students in these classes. Teachers' higher expectations in relation to the potential of classes were observed in classes with higher average socioeconomic status and those without or with only very few students with disabilities. We discuss these findings in light of the mechanisms of TEE.

1. Introduction

Fifty years ago, Rosenthal and Jacobson (1968) released a study that initiated one of the hottest debates in educational psychology. The subject of this ongoing discussion is Teacher Expectancy Effect (TEE), also known as the Pygmalion Effect. TEE fits in with the long tradition of research within the different disciplines of social science that explore how interpersonal expectations create reality (Jussim, Eccles, & Madon, 1996). An impressive body of research has been published since the Pygmalion effect was first described, and these studies contributed to better understanding of this phenomenon. First, a clear distinction between a situation whereby teachers are able to anticipate their students' future outcomes (e.g. their school achievement or social functioning) because they accurately recognize their students' potential, and a situation where distorted expectations produce an effect, has been introduced. The Pygmalion effect applies only to the latter situation (Brophy, 1983; Eccles & Wigfield, 1985). Second, the conditions of biased (inflated or deflated) expectations of teachers were intensively explored. Most of these biases refer to stereotypes associated with the students' race, socioeconomic status (SES), and gender (Ready & Wright, 2011; Rubie-Davies & Peterson, 2016). Third, important data were collected about how teachers communicate their expectations to students and how this translates into student achievement (Brophy, 1983; Harris & Rosenthal, 1985).

However, there are still under-researched aspects of the Pygmalion effect. One of them refers to teacher expectations of the entire school class (Friedrich, Flunger, Nagengast, Jonkmann, & Trautwein, 2015). This paper bridges existing gaps in three ways. First, we examine the effects of teacher expectations on school achievements of students without disabilities in two particular environments: classrooms with and without students with disabilities. We are interested in the TEE on an individual as well on an entire classrooms level. Therefore, we examine whether and to what extent teacher expectations at the beginning of the semester in a new school predict school achievement of their students a year-and-a-half later, when we control for the achievement level on entry. Second, we are interested in the role teacher expectations play in students' self-perception of their own abilities – including their mathematical academic self-concept (MASC) – and the extent to which MASC can mediate the link between teachers' expectations and students' achievements. Finally, we examine the Pygmalion effect in middle schools, while the majority of previous studies was conducted on elementary school students or even preschool children (e.g., Hinnant, O'Brien, & Ghazarian, 2009; Muntoni & Retelsddorf, 2018; Speybroeck et al., 2012).

E-mail addresses: bgszum@poczta.onet.pl (G. Szumski), maciej.karwowski@uwr.edu.pl (M. Karwowski).

[1] Both authors contributed equally to this work.

https://doi.org/10.1016/j.cedpsych.2019.101787

Available online 25 June 2019

2. Teacher expectancy effect

According to the TEE, teachers' expectations lead to students' achievements that are in line with these expectations. Hence, TEE is a form of self-fulfilling prophecy (Brophy & Good, 1974; Rosenthal & Jacobson, 1968). If, therefore, the teacher expects one student to achieve greater progress than another student, the first one tends to actually make greater progress than the other student in the same period of time. The mechanism of how TEE operates is quite complex and includes a sequence of five events (Rosenthal, 1994). First (1), there are certain stimuli that trigger teacher expectations and (2) these stimuli lead to expectations that are (3) communicated to students who (4) change and (5) undertake activities that conform with these expectations, which finally lead to outcomes.

Because TEE depends on many factors that do not necessarily occur in all educational settings, its character is probabilistic: It applies to only some students and takes shape under only certain conditions. Early studies into TEE overestimated its effect size, because these investigations were unable to sufficiently control for teacher expectation accuracy (Jussim & Harber, 2005). For example, the classic study by Rosenthal and Jacobson (1968) rightly assumed that the Pygmalion effect occurs only when teachers' expectations are inadequate. This was so, because this specific study did indeed manipulate the teachers' expectations. In natural settings, however, the effect of TEE is weakened by the fact that teachers generally accurately assess their students' abilities if they have the chance to get to know them (Südkamp, Kaiser, & Möller, 2012). Indeed, the accuracy of teachers expectations increases when they have the chance to collect more information about their students (Dusek & Joseph, 1985). Hence, although nowadays researchers agree that the self-fulfilling prophecy effect does take place, its effect size is visibly lower than the one reported in early works. Meta-analyses indicate that the effect size of the TEE falls between $r = 0.10$ and $r = 0.20$ (Raudenbush, 1984; Rosenthal & Rubin, 1978) or – to use a different metric – applies only to about 5–10% of students (Brophy, 1983). More recent studies conclude that this effect is moderate at best (Kuklinski & Weinstein, 2001; Trusz, 2018).

2.1. Conditions for TEE

As we already highlighted, early research into the Pygmalion effect induced teacher expectations experimentally. Therefore, these studies do not form sufficient evidence that teachers spontaneously formulate biased expectations of student potential. To decide whether the Pygmalion effect occurs in a school's actual life, naturalistic studies where teachers formulate their expectations of their students spontaneously are necessary (Jussim et al., 1996). This type of investigation provides data about actual conditions for school achievement, yet the number of such studies is still relatively low (Trouilloud, Sarrazin, Bressoux, & Bois, 2006; Zhu, Urhahne, & Rubie-Davies, 2018).

Teacher expectations may not predict students' actual potential due to the random and unsystematic character of expectations. In such cases, expectations are irrelevant or imprecise. However, when teachers systematically overvalue or undervalue their students due to a single trait, expectations may be considered imprecise and biased (Ferguson, 2003). Here, then, the Pygmalion effect is applicable mainly to students with particular traits and it leads to favoritism or to disadvantaging certain student groups (Hurwitz, Elliott, & Braden, 2007; Jussim et al., 1996; Rubie-Davies & Peterson, 2016). This bias mainly results from teachers relying on stereotypes associated with students' ethnic or racial origin, gender, and diagnostic label, as well as parents' SES (Hurwitz et al., 2007; Ready & Wright, 2011). Teachers tend to underestimate the potential of students from disadvantaged ethnic and national origin groups (Rubie-Davies, Hattie, & Hamilton, 2006; Tenenbaum & Ruck, 2007). They also tend to underestimate the potential of students from low SES families and overestimate their expectations of middle-class students (Rist, 1970). Moreover, teachers display a tendency to overestimate boys' abilities in math and science (Archambault, Janosz, & Chouinard, 2012), and girls' abilities in languages (Peterson, 2000). Hurwitz et al. (2007) showed that teachers regularly underestimate abilities of students with disabilities.

Also student age serves as a factor that may influence TEE. Interestingly, there are arguments for a conviction that age tends to strengthen TEE as well as for the assumption that older students are in a sense protected from TEE (Jussim, 2012). On the one hand, there are reasons to assume that older students will be more prone to adult (including teachers) expectations than younger students. Indeed, the ability to recognize other people's expectations grows with age (Weinstein, Marshall, Sharp, & Botkin, 1987), as is readiness to include these expectations in one's self-concept – a phenomenon visible in higher concordance between self-rating and teacher-ratings of student abilities (Stipek & Tannatt, 1984; Stipek, 1981). On the other hand, however, it has been also theorized that older students should be less prone to teacher expectations (Good & Nichols, 2001; West & Anderson, 1976), because their self-concept is already more stable (Guay, Marsh, & Boivin, 2003). As the number of studies on TEE on older students is limited and previous findings are equivocal (Kuklinski & Weinstein, 2001; Rosenthal & Jacobson, 1968), additional investigations are necessary to examine the role of student age on TEE.

3. Student academic self-concept as a mediator of TEE

Student behaviors and characteristics that result from teachers' expectations and mediate the link between expectations and school achievements might shed some light on the mechanism of TEE. For example, effectively communicated diverse expectations might influence student academic self-concept (ASC), motivation, as well as involvement and effort expended in learning (Brophy & Good, 1970; Jussim & Eccles, 1992). Links between ASC and effort students put in studying (Jussim & Eccles, 1992), and their involvement in studying are well established (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). This relation may explain why teachers' positive expectations with regard to some students may lead to an increase in their ASC rather than to their greater involvement in studying and consequently, to an increase in their achievement.

ASC is a perception of students' own competences. Much evidence shows that ASC is domain-specific and students are able to assess their own competences in individual domains, such as language, math, or physical fitness quite early (Marsh, Craven, & Debus, 1998). ASC has considerable power to regulate school achievement (Guay et al., 2003; Marsh & Yeung, 1997), yet some studies indicate that this effect is more stable in math than language or natural sciences (Marsh & Yeung, 1997). A meta-analysis demonstrated that the relation between domain-specific ASC and school achievement ranges from $r = 0.49$ to $r = 0.69$ (Möller, Pohlmann, Köller, & Marsh, 2009), though another synthesis reveals a weaker effect ($r = 0.19$ to $r = 0.27$; Huang, 2011). Teacher expectations predict their students' future ASC on both an individual (Marsh et al., 1998; Trouilloud et al., 2006; Trouilloud, Sarrazin, Martinek, & Guillet, 2002) and an entire class level (Rubie-Davies, 2006). Rubie-Davies (2006) showed that teacher expectations with regard to class predict a change in ASC; the decrease due to low expectations greatly exceeds the increase due to high expectations. This is in line with previous findings that the effect of teacher expectations is asymmetric, and that negative expectations are particularly strong (Eccles & Wigfield, 1985).

Finally, there is evidence that ASC might partially mediate the relation between teacher expectations and their students' future school achievement (Friedrich et al., 2015; Kuklinski & Weinstein, 2001; Trouilloud et al., 2002). Trouilloud et al. (2002) observed that the link between teacher expectations and their students' achievement was mediated by the students' perceived ability. It is important to note, though, that ASC was measured concurrently with achievement, so it was not mediation in a strict sense. Moreover, the study lasted for only

10 weeks. Also, Kuklinski and Weinstein (2001) examined whether students' self-expectations mediate TEE. Such mediation was observed only in grade five, but not in grade one or three. Consequently, the effect was not stable. The latest study conducted by Friedrich et al. (2015) showed that ASC mediated the relation between teacher expectations and their students' school grades, but it did not mediate the relation between these expectations and results on the standardized achievement test in math. What is more, the mediation effect was very weak and marginally significant ($b = 0.01$, $p = .05$). Hence, the hypothesis that ACS acts as a mediator between teacher expectations and their students' school achievement requires further investigation.

4. TEE at the class level

The hypothesis that teacher expectations can affect not only individual students, but also the whole class was formulated quite a while ago (Brophy, 1985). However, its theoretical premises are not fully clear. Teachers' expectations of the entire class may depend on group perception rules, because a school class is a specific, formalized social group (Parsons, 2008). Over- or understated teachers' expectations with regards to entire classes, require that these classes distinguish themselves in some respects from other school classes. When classes do not differ in any particular ways from each other, teachers very likely do not formulate any particular expectations of them. A recent study conducted on a large sample of grade one students in German middle schools supports this logic (Friedrich et al., 2015). Although in this study the Pygmalion effect was not observed on the class level, it is important to emphasize that only students from one type of school participated. What's more, the study was conducted immediately after the selection threshold, which is likely to have limited both intra-class and between-classes student ability variance. Similarly, Smith et al. (1998) examined TEE on the level of classes. Students were placed either in homogeneous classrooms that applied intra-class ability grouping or in more heterogenic classes. The authors relied on the labeling hypothesis (Jussim, Palumbo, Chatman, Madon, & Smith, 2000), so they expected a stronger TEE in homogeneous classes. This hypothesis stemmed from the assumption that teachers would overrate student abilities in "gifted" classes and underrate the abilities of students in "low ability" classes. However, the contrary was observed – the Pygmalion effect was found only in heterogeneous classes (Smith et al., 1998). It was posited that grouping may make it easier for teachers to accurately assess their students' abilities, as they gain confidence that all students fall short of or exceed a certain cut-off (Smith et al., 1998). Hence, the label attached to classes that are homogeneous with respect to the level of abilities may increase the accuracy of teachers' recognition of their students' potential, rather than distort their expectations. The situation may be different in classes whose labels do not reflect student abilities. The study conducted by Eden (1990) provides indirect support for this notion. In this study, platoons of school troops were randomly assigned labels that showed either the "high level of potential" label or lack of such a label, and then the effect of self-fulfilling prophecy was tested on a group level. Troops from those platoons whose leaders were informed that their platoons exhibited high potential achieved higher results than did the control group troops (Eden, 1990). This study informs further tests of TEE on a group level – such as school classes whose labels may evoke particular expectations with regards to student achievement. One such class type is an inclusive class that groups together students with and without disabilities.

In many contemporary school systems, the number of inclusive classes is large and keeps growing (Agency, 2012; Ferguson, 2008). Although school achievements of students without disabilities in inclusive classes is higher than in traditional ones (Szumski, Smogorzewska, & Karwowski, 2017), teachers are often concerned that the presence of students with disabilities may inhibit school achievement of students without disabilities (Curcic, 2009). Parents of students without disabilities share these concerns (Martin, 2002; Rafferty, Boettcher, & Griffin, 2001). What's more, in many systems, inclusive classes are localized in districts inhabited by families with lower SES and there is compelling evidence that children from lower SES families have a higher likelihood of being placed in these classes (Demeris, Childs, & Jordan, 2007; Rouse & Florian, 2006). Consequently, schools with a high percentage of students from disadvantaged social groups more frequently educate students with disabilities (Van der Veen, Smeets, & Derriks, 2010), which might lead to teachers forming impressions that inclusive classes are likely to have students with low potential.

5. The present study

The purpose of this study was to examine whether teachers of middle school children create expectations about their students' potential, which then predict those students' achievement in math several months later (i.e., testing the Pygmalion effect on an individual level), and to test whether academic self-concept mediates this effect. We also aimed at examining whether teacher expectations are formulated on the level of school classes and whether the Pygmalion effect occurs in inclusive classes, meaning those classes in which the class context is different due to the presence of students with disabilities. Our hypotheses were three-fold. Firstly (H1), we expected TEE to be mediated by students' academic self-concept. More specifically, we predicted that higher teacher expectations will translate into higher conviction about student competencies and abilities in math (math academic self-concept) that would consequently strengthen their math achievement. Secondly (H2), we theorized that TEE will not only occur on the individual level (i.e. the level of individual students), but also on the level of classes. Hence, we predicted that higher average expectations in relation to the class will strengthen student achievement as well. Thirdly and finally, although or sample included only students without disabilities or special educational needs, we expected (H3) that the number of students with disabilities in the classroom will form a factor negatively linked to teacher expectations on the class level.

The study was run in Polish middle schools located in cities of over 100,000 inhabitants.

Students in grades 7–9 attend middle schools and are aged 13–15 years (EURIDICE, 2008). Middle schools are mandatory and admit elementary school graduates without selection (or tracking). Most often, students attend middle schools in their neighborhood. However, in large cities, some students choose middle schools that have a better reputation. This brings about a relatively high variability of school achievement of students from schools in these cities – in 2012, between-school variance of end-of-middle school examination results exceeded 40% of overall variance (Dolata, Jasińska, & Modzelewski, 2012).

The Polish school system is characterized by a relatively high level of dissemination of inclusive education of students with disabilities (Szumski & Karwowski, 2014, 2015). On the level of middle schools, inclusive education applies to almost 50% of all students with disabilities (GUS, 2016). Each inclusive class usually includes 1–5 students with disabilities. Although these classes are not always specifically labelled as inclusive, they are clearly distinct from other classes in teachers' opinions (DeBoer, Pijl, & Minnaert, 2011). Students with disabilities are diagnosed by multidisciplinary diagnostic teams and they receive the official statement about their disability. This group contains students with intellectual disabilities, autistic spectrum disorders, emotional and behavioral disorders, hearing and visual impairments, physical and multiple disabilities (GUS, 2016). For each students with a statement about disability, teachers, special educators, and school psychologists create individual educational programmes (IEP; see Ryndak et al., 2014, p. 381). These students receive special educators' support of diverse intensity that is adjusted to their needs and grade levels (in and out of class, in the form of specialist therapy – Szumski & Karwowski, 2014).

6. Method

6.1. Participants

The main group of participants in this study was made up of 1488 first-grade middle-school students (M_{age} = 13.15, *SD* = 0.44, 49% females), members of 108 classes attending 40 schools in Poland. The other group of participants were their math teachers (*N* = 97): 82 female (85%) and 15 male (15%). The teachers were between 25 and 64 years old, with *M* = 44.76 and *SD* = 8.50. As the study was longitudinal (see the *Procedure* section), there were some drop-outs between waves and the number of missing values differed across variables (see the *Data Analysis and Imputation of Missing Values* section). As our hypotheses concerned TEE in relation to students without disabilities, only such students were included in our analysis. Thus, we intentionally excluded from all our analyses those students with disabilities (*n* = 211) who attended inclusive classes.

The use of a multilevel and multi-strata selection procedure enabled us to obtain a sample of Polish third-grade students, which might be considered representative for middle schools localized in large cities (excluding special school students and students from very small schools – fewer than 10 students per grade) with an intentional over-representation of inclusive classes. The sample was drawn from the registers of the Polish Educational Information System (PEIS) (https://cie.men.gov.pl/sio-strona-glowna/). Schools were randomly selected and two or three randomly chosen first-grade classes from each school were invited to participate in the study. Out of 108 classes, 34 (31%) had 3 or more students with a disability (min = 3, max = 8). In 28 classes (26%), there were one or two students with a disability and in 46 classes there were no students with a disability. As mentioned above, in all analyses, we rely exclusively on students without disability.[2]

6.2. Measures

6.2.1. Math achievement

To measure math achievement, we conducted a standardized math test. At Time 1 (T1), the test consisted of 16 items (Cronbach's α = 0.89) with varying response formats (e.g. multiple choice items, open questions), while in Time 3 (T3), it consisted of 17 items (Cronbach's α = 0.80). We used a set of 8 fixed items ("anchor items") to be able to compare the test results of T1 and T3 and a set of items that varied between the time points to reduce possible influences of training effects. The test content was based on the school curriculum of middle schools in Poland. Item response theory was used to scale math achievement test scores. We used expected a posteriori person-parameter estimates (EAPs), calculated with XCalibre 4 (Guyer & Thompson, 2011) for further analyses.

6.2.2. Academic self-concept (ASC and MASC)

General ASC was measured in Time 1 using the Short Academic Self-Concept Scale (SASCS) (Szumski & Karwowski, 2015). Sample items included in the SASCS were: (a) "When the teacher explains the new lesson, I understand everything right away," or "I always understand what our teacher explains to us." The scale had high internal consistency (Cronbach's α = 0.85) and its one-factor structure was confirmed by a confirmatory factor analysis (CFA) (*CFI* = 0.988, TLI = 0.976, *SRMR* = 0.02, *RMSEA* = 0.068 [90% CI: 0.048, 0.089]). All items were rated on a four-point scale (1 = *definitely not*, 4 = *definitely yes*). Previous studies demonstrated both convergent and discriminant validity of the SASCS (Karwowski & Szumski, 2015).

Academic self-concept in math (MASC) was measured in Time 2 by the math-related part of the Self-Description Questionnaire-II (Marsh & O'Neill, 1984; Polish adaptation by Karwowski & Szumski, 2015). The items (sample item: "I have always done well in mathematics") were rated on a 7-point scale (*1 = definitely not true of me to 7 = definitely true of me*). The scale had high internal consistency (Cronbach's α = 0.92) and its one-factor structure was confirmed by CFA (*CFI* = 0.982, *TLI* = 0.975, *SRMR* = 0.022, *RMSEA* = 0.069 [90% CI: 0.059, 0.079]).

6.2.3. Teacher expectations

Teachers described all students in the class in terms of a wide variety of characteristics, including seven items that dealt specifically with teachers' perceptions of their students' potential. Sample items included: "(s)he has a very high potential" (see Table A1 in Annex for all items and their descriptive statistics). Teachers described their students on a 5-point Likert scale (*1 = definitely not, 5 = definitely yes*). Each student was rated by his or her math teacher during T1; teacher ratings were modelled as a latent variable. The scale had high internal consistency (Cronbach's α = 0.96) and its one-factor structure was confirmed by CFA (*CFI* = 0.982, *TLI* = 0.969, *SRMR* = 0.017, *RMSEA* = 0.075 [90% CI: 0.062, 0.089]). To create a teacher expectation variable at class level, we aggregated teacher expectations across classes using manifest aggregation (see Marsh et al., 2009).

6.2.4. Control variables

We controlled for student and class socioeconomic status (averaged per class) and the number of students with disabilities per class.

6.2.5. Socioeconomic status (SES)

Consistent with previous research (e.g., Callahan & Eyberg, 2010; Szumski & Karwowski, 2012; Whitley, 2008), the first unrotated factor from a principal components analysis of parental education, number of books in the household, and family possessions was used to create an SES index. The SES factor explained 48% of the latent construct variance and was reliable (α = 0.77).

6.2.6. Number of students with disabilities

For each class, we controlled for the number of students with disabilities per class. It varied from none to 8, with a mean of 1.52 (*SD* = 2.08).

6.3. Procedure

The study was longitudinal, with three measurement waves – the first one (T1) in the beginning of the school year, in the first class of middle school (October); the second one (T2) at the end of the second semester (May); and the third (T3) at the end of third semester (December). Teacher ratings were measured in the first wave together with several teacher and student characteristics that are outside the scope of this study. Student school achievement in math was measured in the first and third wave, while academic self-concept in math was measured in the second wave and their general ASC was measured for control purposes in the first wave. Students solved tests and filled questionnaires in class, in a group setting. Students and parents received general information that the goal of this study is to collect basic data on student functioning in middle schools.

Questions regarding students were included in a questionnaire for the teachers. The remaining questions concerned teachers' education and everyday teaching practices. Teachers were informed that the aim of the study is to describe the specificity of functioning of inclusive classes. Informed consent was obtained from all the teachers and students as well as from school principals and students' parents. The protocol and procedure of the study was accepted by the University of Warsaw, Department of Education Institutional Review Board.

[2] Students with disabilities (*N* = 211) included the following categories: mild intellectual disability (*n* = 60, 28%), students with physical disabilities (*n* = 37, 18%), students with Autism Spectrum Disorder (*n* = 33, 16%), students with emotional or behavioral disorders (*n* = 32, 15%), students with hearing impairments (*n* = 22, 10%), students with multiple disabilities (*n* = 17, 8%), and students with visual disabilities (*n* = 10, 5%).

6.4. Data analysis and imputation of missing values

The data were multilevel: students were clustered in classes, so we relied on multilevel structural equation modelling (SEM) using Mplus 8 (Muthén & Muthén, 2017). The dependent variable was math scores at T3; student math achievement at T1, teacher expectations at T1, and student MASC at T2 served as predictors. In all analyses, we controlled for student SES and for the initial level (T1) of their ASC. Additionally, we included a number of level-2 predictors, specifically: the number of students with disability within a class, averaged SES within a class, and averaged teacher expectations within the class.

All level-1 variables were centered at each class's own mean; level-2 predictors were centered at the sample's grand mean. The number of missing values was moderate: the proportion did not exceed 15% for any of the variables of interest and were missing completely at random: Little (1988) Missing Completely at Random (MCAR) test $\chi^2 = 830.96$, $df = 806$, $p = .264$. We imputed missing data using the multiple imputation option available in Mplus software. Fifty datasets with imputed data were created and then aggregated for the purposes of the analysis. We used the Mplus standard approach to imputation, which is based on a Markov Chain Monte Carlo method, an iterative computational procedure (Enders, 2010).

7. Results

Descriptive statistics and correlations between variables are provided in Table 1. There was a substantial level of stability in math achievement ($r = 0.57$) between T1 and T3. Student SES was robustly related to math achievement at T1 ($r = 0.31$) and T3 ($r = 0.31$). General ASC measured at T1 was related to math achievement at T1 and T3 ($rs = 0.34$ and $r = 0.33$, respectively), MASC measured at T2 ($r = 0.45$) and teacher expectations ($r = 0.41$). MASC measured at T2 was related to math achievement at T1 ($r = 0.31$) and T3 ($r = 0.33$) and teacher expectations ($r = 0.41$). Teacher expectations regarding student potential were linked to student achievement at T1 and T3 (both $rs = 0.42$) (all $ps < 0.001$).

To examine our hypotheses regarding TEE, we proceeded with a structural equation modelling (SEM), regressing student math achievement at T3 on the latent variable of teacher expectations, student SES and their ASC and math achievement at T1, as well as student MASC at T2. Moreover, we included a number of level-2 predictors: Teacher expectations at the class level (averaged across individual student ratings within a class), the averaged SES of a class and a number of students with disabilities within a class (see Fig. 1).

The hypothesized model was characterized by a good fit ($CFI = 0.979$, $TLI = 0.974$, $RMSEA = 0.033$, $SRMR_{within} = 0.030$, $SRMR_{between} = 0.000$) and explained a substantial portion of student math achievement at T3 variability – 32% at individual and 64% at class level. Consistently with our hypothesizing (H1), positive teacher expectations related to individual student potential and translated into higher math achievement at T3 (ß = 0.20, $p < .001$), even when we controlled for math achievement at T1 (ß = 0.24, $p < .001$). At the same time, teachers were quite accurate in assessing their students' potential – correlation between math achievement at T1 and teacher ratings of student potential was robust and significant ($r = 0.37$, $p < .001$). Teacher expectations (ß = 0.24, $p < .001$) and math achievement at T1 (ß = 0.06, $p = .049$) translated into student MASC at T2 (even when T1 ASC was controlled for), which then predicted student math achievement at T3 (ß = 0.10, $p < .01$).

Among level-2 predictors, consistent with H2, teacher expectations on the level of class were positively related to math achievement at T3 (ß = 0.40, $p < .001$) and average SES in the classrooms (ß = 0.51, $p < .001$). The number of students with disabilities in the classroom was unrelated to the math achievement of their peers without disability at T3 (ß = −0.03, $p = .66$), when teacher expectations and average SES were controlled, but the classes with a smaller number of students with disabilities had higher average SES (ß = −0.19, $p = .02$). Also, there was a negative effect of the prevalence of students with disabilities in a class on teacher class-average expectations (ß = −0.20, $p = .043$), consistently with H3. Class average SES translated significantly into teacher expectations regarding the potential of the class (ß = 0.45, $p < .001$).

To examine whether student MASC mediates the relationship between teacher expectations and student math achievement at T3 (as assumed by H1), we estimated indirect effects. More specifically, we applied a bootstrap-corrected method to estimate indirect effects of teacher expectations on student achievement at T3 via student MASC. Additionally, we also estimated a potential link between math achievement at T1 and T3 as mediated by MASC, as well as the mediating path: ASC → MASC → math achievement at T3. A similar procedure was conducted on between-level variables; we estimated indirect effects of the number of students with disabilities in mediating the link between average SES in class and teacher expectations, as well as three mediated effects related to student math achievement at T3: (1) average SES → number of students with disability → teacher expectations (class-average) → math achievement at T3, (2) average SES → teacher

Table 1
Descriptive statistics and correlations between observed variables.

		Min	Max	M	SD	2	3	4	5	6
	Level-1									
1	Mat1	−1.45	2.44	0.11	0.89	0.57***	0.34***	0.31***	0.42***	0.31***
2	Mat3	−1.12	2.61	0.35	0.94		0.33***	0.33***	0.42***	0.31***
3	ASC	1	4	2.62	0.69			0.45***	0.41***	0.22***
4	MASC	1	7	4.37	1.54				0.41***	0.11**
5	T-Exp	1	5	3.47	0.95					0.22***
6	SES	−2.8	2.26	0.03	1					
	Level-2									
1	AveSES	−1.59	1.64	−0.05	0.62	0.47***	−0.19$^{\#}$			
2	AveT-Exp	8.44	16.21	12.68	1.6		−0.25*			
3	Disab$^{\#}$	0	8	1.52	2.08					

Note. $N = 1488$ for variables measured at Level 1 and $N = 108$ for variables measured at Level-2.
Mat1 = Math Achievement at T1, Mat3 = Match Achievement at T3, #Disab = Number of Students with Disability in Class, SES = Socioeconomic Status, SES-Class = Socioeconomic Status averaged across classes; TEE-Class = Teachers' Expectations at the Class Level, TEE-Indiv = Teacher Expectations at the Individual Level; ASC = Academic Self-Concept in Math.

# $p < .06$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

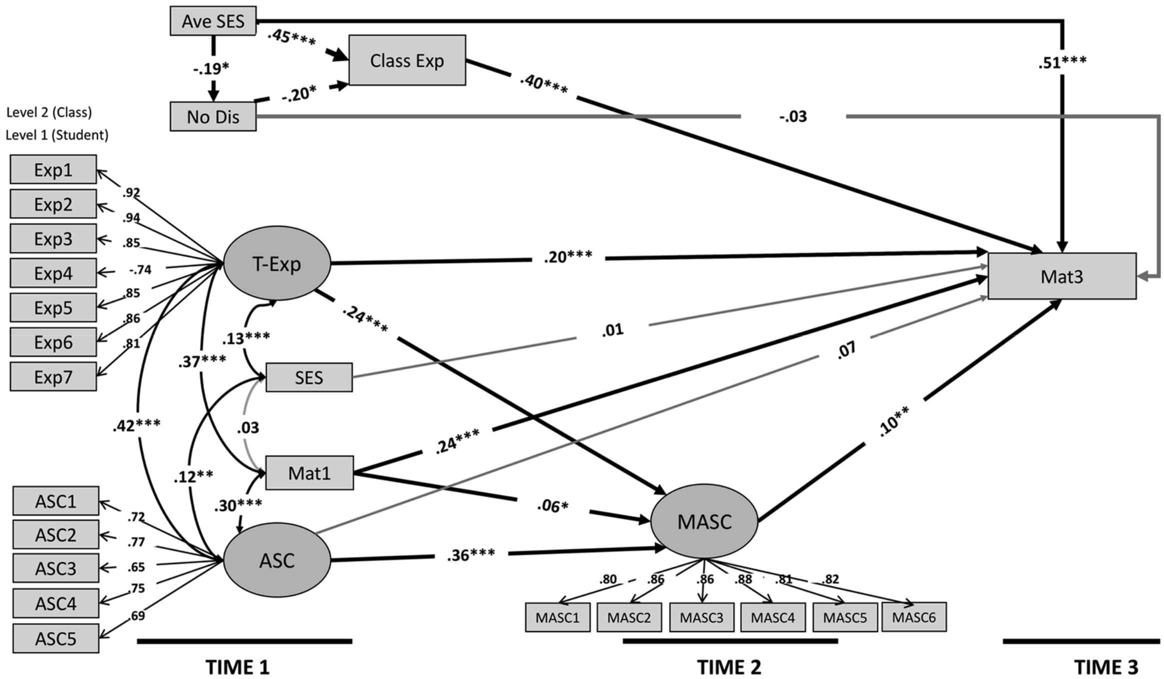

Fig. 1. Multilevel longitudinal structural equation model. *Note.* Ave SES = SES averaged within classes; No Dis = Number of Students with Disabilities within a class; SES = socioeconomic status; Exp1-Exp7 = Teacher expectancy items (see appendix for a list and descriptive statistics), Mat1 = math achievement at T1; Mat3 = math achievement at T3; ASC = academic self-concept (general); MASC = math academic self-concept; ASC1-ASC5 = academic self-concept items; MASC1-MASC8 = math academic self-concept items; T-Exp = teachers' expectations latent variable; Class Exp = teacher expectation toward the whole class (averaged). Only direct effects are presented; see Table 2 for indirect effects.

Table 2
The summary of indirect effects observed in this study.

Indirect effects	Standardized estimate	*SE*	95% *BC CI*	*p*
Level-1				
T-Exp → MASC → MAT3	0.024	0.008	0.007–0.040	0.005
Mat1 → MASC → MAT3	0.006	0.004	−0.002–0.014	0.127
ASC → MASC → MAT3	0.036	0.014	0.009–0.062	0.009
Level-2				
AveSES → #Disab → Class-Exp	0.037	0.024	−0.010–0.084	0.121
AveSES → #Disab → Class-Exp → MAT3	0.015	0.010	−0.005–0.035	0.149
AveSES → Class-Exp → MAT3	0.181	0.049	0.085–0.278	< 0.001
#Disab → Class-Exp → MAT3	−0.078	0.043	−0.162–0.006	0.068

Note. T-Exp = Teacher Expectations, MASC = Academic Self-Concept in Math, ASC = Academic Self-Concept, Mat1 = math achievement at T1; Mat3 = math achievement at T3, AveSES = Socioeconomic Status Averaged per Class, Class Exp = teacher expectation toward the whole class (averaged), #Disab = Number of Students with Disability in Class, *SE* = Standard Error, 95% *BC CI*: 95% bootstrap corrected confidence intervals.

expectations (class-average) → math achievement at T3 and (3) number of students with disability → teacher expectations (class-average) → math achievement at T3 (Table 2).

As illustrated in Table 2, the main indirect effect of interest, i.e. the link between teacher expectations and student math achievement at T3 as mediated by student MASC was statistically significant (standardized estimate = 0.024, p = .005, 95% CI: 0.007–0.040). This is similar to the indirect effect of ASC and math achievement at T3 as mediated by student MASC at T2 (standardized estimate = 0.036, p = .009, 95% CI: 0.009–0.062). The indirect effect of T1 and T3 math achievement as mediated by T2 MASC did not reach the conventional level of statistical significance (standardized estimate = 0.006, p = .13, 95% CI: −0.002–0.014).

Among between-level indirect effects, only the relationship between class average SES and math achievement at T3 as mediated by teacher expectations was statistically significant (standardized estimate = 0.18, p < .001, 95% CI: 0.085–0.278); the number of students with disabilities did not mediate the link between average SES and teacher expectations (standardized estimate = 0.037, p = .12, 95% CI: −0.010–0.084), although there were marginal negative indirect effects of the relationship between the number of students with disabilities in class and math achievement at T3, as mediated by teacher expectations toward class (standardized estimate = −0.078, p = .068, 95% CI: −0.162–0.006) (see Table 2).

8. Discussion

This study had three basic purposes. The first one was to examine whether teachers of Polish middle school children create unreasonable expectations when it comes to their students' achievements, which then predict these achievements in math after 1.5 years of education: the Pygmalion effect on the level of an individual. Our second goal was to examine whether MASC mediates the Pygmalion effect. Thirdly and finally, we examined whether teacher expectations are also formulated on the level of school classes, meaning whether the Pygmalion effect occurs in inclusive classes.

8.1. The Pygmalion effect on the individual level and the mediating role of MASC

The results we obtained confirm the occurrence of the Pygmalion Effect on an individual level. The overall effect size of teacher expectations was moderate (ß = 0.20): stronger than the effect observed in previous Pygmalion investigations (Raudenbush, 1984; Rosenthal & Rubin, 1978), yet similar when it comes to math achievement (Jussim & Eccles, 1992; Jussim, 1989). Our study shows that the Pygmalion effect is present not only among elementary school students, but also in middle schools. What's more, the effect size of this phenomenon is comparable with that observed among younger students (Shenke, Nguyen, Watts, Sarama, & Clements, 2017; Speybroeck et al., 2012). Thus, our findings go against the hypothesis that adolescents are more resistant to the biased expectations of their teachers (e.g., Good & Nichols, 2001; Jussim, 2012; West & Anderson, 1976). This modest (yet still robust) effect size could be caused by the fact that this study was conducted in middle schools – the educational stage that is characterized by growing importance of school achievements and competition between students (e.g., Simmons, 2017). Emphasis on both achievement and competition are risk-factors for stronger TEE (Good & Brophy, 2000). Going further, as TEE is usually disadvantageous, because negative effects of underexpectations are usually stronger than positive effects of overexpectations (Jussim et al., 1996), TEE requires well-thought interventions also in middle and high schools. The observed link between teacher expectations and student achievements in math occurred when initial achievement, ASC, and the SES of their parents were controlled. SES is the factor that was rightly regarded as a key variable that distorts the expectations of teachers in relation to the educational capabilities of their students (Rist, 1970; Rubie-Davies & Peterson, 2016), but also an important predictor of school achievements (Sirin, 2005). In the initial measurement, we also controlled teacher accuracy in recognizing student potential: a standard solution in contemporary studies of the Pygmalion effect (Jussim & Harber, 2005). The correlation between teacher expectations and their students' initial achievements was robust ($r = 0.37$), which confirms the relatively high level of accuracy of teacher assessments (Südkamp et al., 2012). Hence, this study provides further support to the argument that the accuracy of teacher assessments is stronger than the TEE (Jussim & Harber, 2005) even at the beginning of a new stage of education, i.e. when teachers do not have a sufficient amount of time to become acquainted with student abilities.

So far, only a small number of studies examined the academic self-concept as a mediator of the Pygmalion effect. Moreover, these studies brought about ambiguous results (Friedrich et al., 2015; Kuklinski & Weinstein, 2001; Trouilloud et al., 2002). Our study showed that indeed domain-specific MASC partially mediates the effect of teacher expectations with regards to school achievements in math. Numerous studies previously showed that teachers use many such ways of communicating their expectations to students that may influence their motivational characteristics, including their self-concept. These are, among others: more frequently giving praise to students deemed more promising, accepting their ideas, maintaining visual contact with them, or smiling at them (Brophy, 1983; Harris & Rosenthal, 1985). Previous studies also showed that ASC predicts student school achievements (Guay et al., 2003), and this effect is especially strong with respect to achievements in math (Marsh & Yeung, 1997).

Our results extend findings from a recent study conducted by Friedrich et al. (2015), which showed that student MASC mediates the Pygmalion effect when school grades rather than test results are used as a measure of achievement in math. The difference between this study and the Friedrich et al. (2015) investigation may result from contextual characteristics. Even though both studies were conducted in middle schools, Germany has tracking on this school level, whereas Poland does not. The study by Friedrich and colleagues was conducted only within the lowest school track, where the variance of student potential may be restricted in comparison with the Polish middle school. Grouping may help teachers to accurately perceive their student abilities, because it decreases student variability and provides the teacher with certainty that all are below or above a certain cut-off (Smith et al., 1998). In heterogeneous classes, teachers may find it more difficult to properly assess individual students, and consequently a greater probability that the TEE will actually occur does exist. In fact, both the Pygmalion effect and the influence of teacher expectations on MASC in the German study were relatively weak, while this study found them stronger.

The mediation we observed is not a full mediation, which indicates that other mediators of the relation between teacher expectations and student school achievements might exist. These mediators were not included in this study and may take the form of other motivational variables (such as self-efficacy that according to many researchers makes it possible to more accurately predict school achievements than the ASC does, (Bong & Skaalvik, 2003; Pajares & Schunk, 2001), but more importantly of instructional strategies that differentiate the process of student learning. Many behaviors of teachers that result from their expectations exert more influence on their students' cognitive processes than on their motivational sphere. Teachers give more and more difficult tasks to high-expectancy students, create more opportunities to answer, provide more frequent and more accurate guidance, and ask more stimulating questions. All of these elements make the process of these students' learning more effective and lead to higher achievements (Harris & Rosenthal, 1985). It is important to emphasize, though, that previous studies focused on the role of teachers' behaviors that result from their expectations rather than on how students' traits and behaviors may result from teachers' convictions (Harris & Rosenthal, 1985). Further studies of the mediating role of students' cognitive and motivational characteristics are necessary to fully untangle the mechanism of the influence of teacher expectations on student school achievements. Last, but not least, we emphasize that the practical importance of the indirect effect of teacher expectations on student achievements as mediated by students' math academic self-concept should not be overestimated, given the tiny effect size we observed (ß = 0.024) and keeping in mind that previous studies (e.g. Friedrich et al., 2015) obtained even weaker indirect effects (ß = 0.01). Thus, future studies should focus on examining other theoretically derived potential mediating effects.

8.2. The Pygmalion effect at classroom level

The main purpose of this study was to ascertain whether the Pygmalion effect occurs in inclusive classrooms when it comes to students without disabilities. The pattern of findings we obtained is consistent with our hypotheses. Teacher expectations with regards to their classes were related to average classroom SES and to the number of students with disabilities in the given classroom. Higher average SES in a class resulted in higher expectations, while more students with disabilities led to lower expectations. While the observed relation between SES and teacher expectations on the level of classrooms replicates previous results (Ready & Wright, 2011), the relation between the number of students with disabilities in the classroom and teacher expectations of the potential of students without disabilities forms a new finding. Previous studies showed that teachers do not appreciate the potential of students with disabilities, but scholars hypothesized that this effect stems from an overestimation of the potential of their peers without disabilities, known as the contrast effect (Hurwitz et al., 2007). Rather, our study suggests the assimilation effect, i.e. teachers transfer their lower expectations of students with disabilities onto their peers from the same classroom – an effect that remained statistically significant even when we controlled for the differences in average SES in classrooms. Therefore, we may conclude that lower expectations toward classes with a higher number of students with disabilities were not merely an epiphenomenon of lower SES in these classrooms. Although indeed an average classroom SES played significant role in shaping teacher expectations, accumulation of students with disabilities in the classroom clearly lowered teacher expectations also toward students without disability attending the same class. This phenomenon resembles the courtesy stigma phenomenon, convincingly documented in literature (Ali, Hassiotis, Strydom, & King, 2012). Consistent with the courtesy stigma effect, stereotypes or prejudices with regards to labelled individuals also apply to those who are closely linked to them (Birenbaum, 1992). Importantly, prevalence of students with lower SES in inclusive classes was previously observed in investigations conducted in other school systems (see e.g., Demeris et al., 2007; Rouse & Florian, 2006). The impact of classroom SES and the number of students with disabilities on teacher expectations toward students without disabilities should be of particular concern. Inclusive education is often defined as a new model of education that responds to the diverse needs of modern societies (Artiles, Kozleski, Dorn, & Christstensen, 2006; Opertti, Walker, & Zhang, 2014) and provides optimal educational conditions for all students (Booth & Ainscow, 2002). Its implementation, however, often does not lead to a general improvement in educational practices of schools and teachers, but only to providing support for students with disabilities (O'Rourke, 2015). There is, therefore, a need to enrich teacher education programmes that reduce their biases towards students with low SES, ethnic and linguistic minorities, and students with disabilities, and enhance their ability to work in differentiated classes. Such programmes can contribute to reducing biases against certain groups of students (Kumar & Hamer, 2013; López, 2017).

8.3. Limitations and future studies

Even though this study was conducted on a large sample and utilized a longitudinal design and latent variables analyses, it is not void of limitations that should be addressed in future endeavors. Among the main shortcomings of the current investigation, we perceive the lack of inclusion of other mediators that could explain the mechanisms that stand behind TEE. Consistent with the classic Rosenthal (1994), specific instructional strategies and teachers' behaviors should be incorporated as additional mediators in future longitudinal investigations. Future researchers should also aim at more dynamic and naturalistic measurements of teacher behaviors that are directed toward specific students. This would allow scholars to go beyond self-report data. Including additional variables measured on an individual student's level might be informative as well. We were able to demonstrate that teacher expectations are related to student MASC, even when the initial level of their academic self-concept was controlled. It is important to emphasize, however, that we measured general rather than math self-concept in the first wave of our study. Consequently, future studies should replicate our findings using the same measures of the variables of interests.

Although strong causal statements are unwarranted given our design, there are indeed plausible reasons to believe that teacher expectations and specific behaviors influence students' self-perception. It would be informative to examine whether these practices are also able to build students' self-efficacy and mastery- or avoidance-oriented motivation or other characteristics predicting their math achievement, whether it be low anxiety (Meece, Wigfield, & Eccles, 1990), lack of learned helplessness (Krejtz & Nezlek, 2016), or other factors.

Finally, the specific sociocultural context of our study could influence the overall pattern of the findings as well as more specific effect sizes of our main relationships of interests. Indeed, future meta-analyses and cross-cultural studies should pay more attention to plausible sociocultural moderators of the expectancy effects.

9. Conclusion

Although teachers are quite accurate while assessing their students' potential, their accuracy is far from perfect. This study demonstrated that teacher expectations related to individual student as well as more general class potential were linked to student math achievement one-and-a-half-years later, even if the initial level of math achievement was statistically controlled for. The observed indirect effect demonstrated that teacher expectations were related to students' self-perception – i.e., their MASC, which then translated into achievement. Additionally, the Pygmalion effect was observed not only on the individual student level, but also at class level. In other words, the higher teacher expectations are with regard to the whole class, the higher the achievements of individual students will be.

Acknowledgement

This study was supported by Contract No. UMO-2012/07/B/HS6/01442 from the National Science Centre Poland.

Appendix A

Table A1
Descriptive statistics and correlations between teacher expectations items.

	Items	M	SD	1	2	3	4	5	6	7
1	(S)he has a very high potential	3.47	1.11	1						
2	(S)he is more promising compared to her/his peers	3.47	1.07	0.87	1					
3	I think (s)he'll grow a lot this year	3.55	0.97	0.74	0.79	1				
4	(S)he thinks slower than her/his classmates	2.34	1.25	−0.64	−0.65	−0.56	1			
5	(S)he has great potential	3.29	1.11	0.77	0.77	0.76	−0.54	1		
6	(S)he quickly learns new things]	3.43	1.08	0.83	0.77	0.70	−0.61	0.74	1	
7	(S)he is more astute than her/his classmates	3.35	1.11	0.74	0.73	0.64	−0.62	0.67	0.80	1

Note. $N = 1488$, all correlations are statistically significant ($p < .001$).

References

Ali, A., Hassiotis, A., Strydom, A., & King, M. (2012). Self-stigma in people with intellectual disabilities and courtesy stigma in family careers: A systematic review. *Research in Developmental Disabilities, 33*, 2122–2140.
Archambault, I., Janosz, M., & Chouinard, R. (2012). Teacher beliefs as predictors of adolescents' cognitive engagement and achievement in mathematics. *Journal of Educational Research, 105*, 319–328.
Artiles, A. J., Kozleski, E., Dorn, S., & Christensen, C. (2006). Learning in inclusive education research: Re-mediating theory and methods with a transformative agenda. *Review of Research in Education, 30*, 65–108.
Birenbaum, A. (1992). Courtesy stigma revisited. *Mental Retardation, 30*, 265–268.
Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1–40.
Booth, T., & Ainscow, M. (2002). *Index for inclusion: Developing learning and participation in*

schools. Bristol: Centre for Studies on Inclusive Education.
Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75*, 631–661.
Brophy, J. (1985). Teacher-student interaction. In J. B. Dusek (Ed.). *Teacher expectancies* (pp. 303–328). Hillsdale, NY: Erlbaum.
Brophy, J., & Good, T. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*, 365–374.
Brophy, J., & Good, T. (1974). *Teacher-student relationships: Causes and consequences.* New York: Holt, Rinehart and Winston.
Callahan, C. L., & Eyberg, S. M. (2010). Relations between parenting behavior and SES in a clinical sample: Validity of SES measures. *Child & Family Behavior Therapy, 32*, 125–138.
Curcic, S. (2009). Inclusion in PK-12: An international perspective. *International Journal of Inclusive Education, 13*, 517–538.
DeBoer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary school teachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education, 15*, 331–353.
Demeris, H., Childs, R. A., & Jordan, A. (2007). The influence of students with special needs included in grade-3 classrooms on the large-scale achievement scores of students without special needs. *Canadian Journal of Education, 30*, 609–627.
Dolata, R., Jasińska, A., & Modzelewski, M. (2012). *Wykorzystanie krajowych egzaminów jako instrumentu polityki oświatowej na przykładzie różnicowania się gimnazjów w dużych miastach.* [Using nationwide exams as instruments of educational policy on an example of differentiation of middle-schools in large cities]. *Polityka Społeczna, 1*, 41–46.
Dusek, J. B., & Joseph, G. (1985). The bases of teacher expectancies. In J. B. Dusek (Ed.). *Teacher expectancies* (pp. 229–250). Hillsdale, NY: Erlbaum.
European Agency for Development in Special Needs Education, EADSNE. (2012). Special needs education. Country date 2012. Odense: European Agency for Development in Special Needs Education.
Eccles, J., & Wigfield, A. (1985). Teacher expectations and student motivation. In J. Dusek (Ed.). *Teacher expectancies* (pp. 185–226). Hillsdale, NY: Erlbaum.
Eden, D. (1990). Pygmalion without interpersonal contrast effects: Whole groups gain from raising manager expectations. *Journal of Applied Psychology, 75*, 394–398.
Enders, C. K. (2010). *Applied missing data analysis.* NY, London: The Guilford Press.
EURIDICE (2008). *The system of education in Poland.* Warsaw: Foundation for the Development of the Education System http://eurydice.org.pl/wpcontent/uploads/2014/10/the_system_2008.pdf.
Ferguson, R. F. (2003). Teachers' perceptions and expectations and the Black-White test score gap. *Urban Education, 38*, 460–507.
Ferguson, D. L. (2008). International trends in inclusive education: The continuing challenge to teach each one and everyone. *European Journal of Special Needs Education, 23*, 109–120.
Friedrich, A., Flunger, B., Nagengast, B., Jonkmann, K., & Trautwein, U. (2015). Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement. *Contemporary Educational Psychology, 41*, 1–12.
Good, T. L., & Brophy, J. E. (2000). *Looking in classrooms* (8th ed.). New York: Longman.
Good, T. L., & Nichols, S. L. (2001). Expectancy effects in the classroom: A special focus on improving the reading performance of minority students in first-grade classrooms. *Educational Psychologist, 36*, 113–126.
Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology, 95*, 124–136.
GUS (2016). *Oświata i wychowanie w roku szkolnym 2015/2016 [Education in 2015/2016 school year].* Warsaw: Zakład Wydawnictw Statystycznych.
Guyer, R., & Thompson, N. (2011). *Item response theory parameter recovery using XcalibreTM 4.1. Technical Report.* Sain Paul, MN: Assessment System Corporation.
Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin, 97*, 363–386.
Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology, 101*, 662–670.
Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*, 505–528.
Hurwitz, J. T., Elliott, S. N., & Braden, J. P. (2007). The influence of test familiarity and student disability status upon teachers' judgments of students' test performance. *School Psychology Quarterly, 22*, 115–144.
Jussim, L. (2012). *Social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy.* New York: Oxford University Press.
Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology, 57*, 469–480.
Jussim, L., & Eccles, J. S. (1992). Teacher expectations II: Construction and reflection of student achievement. *Journal of Personality and Social Psychology, 63*, 947–961.
Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology, 28*, 281–388.
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.
Jussim, L., Palumbo, P., Chatman, C., Madon, S., & Smith, A. (2000). Stigma and self-fulfilling prophecies. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.). *The social psychology of stigma* (pp. 374–418). New York: Guilford Press.
Karwowski, M., & Szumski, G. (2015). Self-description questionnaire: Polish adaptation. Unpublished manuscript.
Krejtz, I., & Nezlek, J. B. (2016). It's Greek to me: Domain specific relationships between intellectual helplessness and academic performance. *Journal of Social Psychology, 156*, 664–668.
Kuklinski, M. R., & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development, 72*, 1554–1578.
Kumar, R., & Hamer, L. (2013). Preservice teachers' attitudes and beliefs toward student diversity and proposed instructional practices: A sequential design study. *Journal of Teacher Education, 64*(2), 162–177.
Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198–1202.
López, F. A. (2017). Altering the trajectory of the self-fulfilling prophecy: Asset-based pedagogy and classroom dynamics. *Journal of Teacher Education, 68*, 193–212.
Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69*, 1030–1053.
Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*, 764–802.
Marsh, H. W., & O'Neill, R. (1984). Self-Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*, 153–174.
Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397–416.
Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*, 41–54.
Martin, C. (2002). Serving gifted students trough inclusion: A parent's perspective. *Roeper Review, 24*, 127.
Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*, 60–70.
Möller, J., Pohlmann, B., Köller, O., & Marsh, H. (2009). A meta-analytic path analysis of the Internal/External Frame of Reference Model of academic achievement and academic self-concept. *American Educational Research Journal, 48*, 1315–1346.
Muntoni, F., & Retelsddorf, J. (2018). Gender-specific teacher expectations in reading – The role of teachers' gender stereotypes. *Contemporary Educational Psychology, 54*, 212–220.
Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
O'Rourke, J. (2015). Inclusive schooling: If it's so good – why is it so hard to sell? *International Journal of Inclusive Education, 19*, 530–546.
Opertti, R., Walker, Z., & Zhang, Y. (2014). Inclusive education: From targeting groups and schools to achieving quality education as the core of EFFA. In L. Florian (Ed.). *The Sage handbook of special education* (pp. 149–169). Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.
Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding, & S. Rayner (Eds.). *Perception* (pp. 239–266). London: Ablex Publishing.
Parsons, T. (2008). The school class as a social system. In J. H. Ballantine, & J. Z. Spade (Eds.). *Schools and society. A sociological approach to education* (pp. 80–85). (3rd ed.). Los Angeles, London, New Delhi, Singapore: Pine Forge Press.
Peterson, S. (2000). Grades four and eight students' and teachers' perceptions of girls' and boys' writing competencies. *Reading Horizons, 40*, 253–271.
Rafferty, Y., Boettcher, C., & Griffin, K. W. (2001). Benefits and risks of reverse inclusion for preschoolers with and without disabilities: Parents' perspectives. *Journal of Early Intervention, 24*, 266–286.
Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology, 76*, 85–97.
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*, 335–360.
Rist, R. C. (1970). Student social class and teacher expectations: The self–fulfilling prophecy in ghetto education. *Harvard Educational Review, 40*, 411–451.
Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3*, 176–179.
Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart & Winston.
Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences, 3*, 377–386.
Rouse, L., & Florian, L. (2006). Inclusion and achievement: Student achievement in secondary schools with higher and lower proportions of pupils designated as having special educational needs. *International Journal of Inclusive Education, 10*, 481–493.
Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: Exploring relationships. *Psychology in the Schools, 43*, 537–552.
Rubie-Davies, C. M., Hattie, J., & Hamilton, R. (2006). Expecting the best for students: Teacher expectations and academic outcomes. *British Journal of Educational Psychology, 76*, 429–444.
Rubie-Davies, C. M., & Peterson, E. R. (2016). Relations between teachers' achievement, over- and underestimation, and students' beliefs for Māori and Pākehā students. *Contemporary Educational Psychology, 4*, 72–83.
Ryndak, D. L., Reed, D. S., Szumski, G., Orlando, A.-M., Smogorzewska, J., & Gao, W. (2014). Societal inclusion and equity internationally: Initiatives, illustrations, challenges, and recommendations. In M. Agran, F. Brown, C. Hughes, C. Quirk, & D. L. Ryndak (Eds.). *Equity & full participation for individuals with severe disabilities. A vision*

for the future (pp. 371–394). Baltimore, London, Sydney: Paul H. Brookes.
Shenke, K., Nguyen, T., Watts, T. W., Sarama, J. H., & Clements, D. H. (2017). Differential effects of the classroom on African American and non-African American's mathematics achievement. *Journal of Educational Psychology, 109*, 794–811.
Simmons, R. G. (2017). *Moving into adolescence: The impact of pubertal change and school context.* London: Routledge.
Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417–453.
Smith, A. E., Jussim, L., Eccles, J., VanNoy, M., Madon, S., & Palumbo, P. (1998). Self-fulfilling prophecies, perceptual biases, and accuracy at the individual and group level. *Journal of Experimental Social Psychology, 34*, 530–561.
Speybroeck, S., Kuppens, S., Van Damme, J., Van Petegem, P., Lamote, C., Boonen, T., & de Bilde, J. (2012). The role of teachers' expectations in the association between children's SES and performance in kindergarten: A moderated mediation analysis. *PLoS One, 7*(4), e34502.
Stipek, D. J. (1981). Children's perceptions of their own and their classmaters' ability. *Journal of Educational Psychology, 73*, 404–410.
Stipek, D. J., & Tannatt, L. M. (1984). Children's judgements of their own and their peers' academic competence. *Journal of Educational Psychology, 76*, 75–84.
Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*, 743–762.
Szumski, G., & Karwowski, M. (2012). School achievement of children with intellectual disability: The role of socioeconomic status, placement, and parents' engagement. *Research in Developmental Disabilities, 33*, 1615–1625.
Szumski, G., & Karwowski, M. (2014). Psychosocial functioning and school achievement of children with mild intellectual disability in Polish special, integrative, and mainstream schools. *Journal of Policy and Practice in Intellectual Disabilities, 11*, 99–108.
Szumski, G., & Karwowski, M. (2015). Emotional and social integration and the big-fish-little-pond effect among students with and without disabilities. *Learning and Individual Differences, 43*, 63–74.
Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. *Educational Research Review, 21*, 33–54.
Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology, 99*, 253–273.
Trouilloud, D., Sarrazin, P., Bressoux, P., & Bois, J. (2006). Relation between teachers' early expectations and students' later perceived competence in physical education classes: Autonomy-supportive climate as a moderator. *Journal of Educational Psychology, 98*, 75–86.
Trouilloud, D. O., Sarrazin, P. G., Martinek, T. J., & Guillet, E. (2002). The influence of teacher expectations on student achievement in physical education classes: Pygmalion revisited. *European Journal of Social Psychology, 32*, 501–607.
Trusz, S. (2018). Four mediation models of teacher expectancy effects on students' outcomes in mathematics and literacy. *Social Psychology of Education, 21*, 257–287.
Van der Veen, I., Smeets, E., & Derriks, M. (2010). Children with special educational needs in the Netherlands: Number, characteristics and school career. *Educational Research, 52*, 15–43.
Weinstein, R. S., Marshall, H. H., Sharp, L., & Botkin, M. (1987). Pygmalion and the student: Age and classroom differences in children's awareness of teacher expectations. *Child Development, 58*, 1079–1093.
West, C., & Anderson, T. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research, 46*, 185–213.
Whitley, J. (2008). A model of general self-concept for students with learning disabilities: Does class placement play a role? *Developmental Disabilities Bulletin, 36*, 106–134.
Zhu, M., Urhahne, D., & Rubie-Davies, Ch. M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology, 38*, 648–668.