The Effect of Pullout Classes on Gifted Students in Primary Education: A Systematic Literature Review

Elzelien Hopma Zijlema (S4940644)

Faculty of Behavioral and Social Sciences, University of Groningen

PABA-A412: Bachelor Project

Dr. Run Tan

06-06-2025

Samenvatting

Steeds meer landen maken de beweging naar inclusief onderwijs. Er zijn twijfels of dit goed is voor hoogbegaafde kinderen, omdat in de klas de focus vaak ligt op kinderen met andere ondersteuningsbehoeften. Onderwijs aan hoogbegaafde kinderen wordt daarom juist steeds meer apart gegeven, met name in de vorm van plusklassen. Er is echter onvoldoende onderzoek naar het effect hiervan. Dit leidde tot de vraag: wat is het effect van plusklassen als onderdeel van regulier onderwijs op de academische resultaten van hoogbegaafde kinderen in het primair onderwijs? Om deze vraag te beantwoorden is een systematische literatuurreview uitgevoerd, gebaseerd op de PRISMA-richtlijnen. Dit heeft zeven onderzoeken naar deeltijd hoogbegaafdheidsonderwijs opgeleverd, variërend in opbouw en uitvoer. De onderzoeken laten positieve effecten zien, zowel op academische als affectieve tests. Mogelijke factoren die aan de positieve effecten bijdragen zijn een breed curriculum, individuele aandacht, motivatie en ondersteuning van de familie. Verder lijken goede programmadoelen met theoretische onderbouwing te zorgen voor betere programma-effecten. Ondanks dat het hier gaat om een kwalitatieve analyse, lijkt dit onderzoek in lijn met de eerdere meta-analyse van Vaughn (1991). Deeltijdklassen lijken dan ook een passende vorm van onderwijs voor hoogbegaafde kinderen. Vooruitkijkend is er behoefte aan evaluatie en onderzoek van de in de praktijk gebruikte programma's, en de implementatie van al onderzochte programma's in de praktijk.

Keywords: Gifted Students, Gifted education, Pullout programs, Academic Achievement

Abstract

In the last thirty years, fundamental policy shifts have been made towards inclusive education. There are doubts about whether inclusive education is the right fit for gifted students. Thus, gifted education has become more separate during this period, often in the form of pullout classes. However, there is a lot of ambiguity regarding effective practices. This raises the question: *What is the effect of pullout classes as part of mainstream education on the academic performance of gifted students in primary school?* To answer this question, a systematic literature review was conducted, based on the PRISMA guidelines. This yielded seven articles about pullout programs, varying in both design and implementation. The included studies show a positive effect, both on academic and non-academic measures. Possible factors contributing to the positive effects of the pullout programs are the use of a rich curriculum, individualized attention, increased motivation, and changed family support. Furthermore, good program goals with strong theoretical links and program fidelity seem to improve the effects of a pullout program. Despite this being a qualitative analysis, this research seems to be in line with Vaughn's (1991) earlier meta-analysis. Therefore, part-time classes appear to be an appropriate form of education for gifted children. Moving forward, there is a need for the assessment of the broad programs currently used and the implementation of research-based programs.

Keywords: Gifted Students, Gifted education, Pullout programs, Academic Achievement

According to the US Supreme Court, “separate is inherently unequal” *(Brown v. Board of Education,* 1954). This laid the groundwork for inclusive education in the US, as this led to the inclusion of all students into the regular classroom (Yell, 2021). Inclusive education means that teachers create activities to make them available to all students (van Gerven, 2021). By integrating students with disabilities into regular classrooms, all students have the same educational opportunities (Jardinez & Natividad, 2024).

Inclusive education has been a topic of discussion for quite some time. As a result of this, in the last 30 years, fundamental policy shifts have been made towards inclusive education. To start, the Salamanca Statement of 1994 declares that students with special educational needs must have access to regular schools (UNESCO, 1994). The need for inclusive education is repeated in Article 24 of the “Convention on the Rights of Persons with Disabilities”, based on the notion of equal opportunity (United Nations, 2006). Furthermore, in 2015, the UN’s Sustainable Development Goal 4, Quality Education, described equal access to education for people with disabilities as a goal to be reached in 2030 (United Nations, 2020).

Inclusive education offers several benefits, including creating acceptance among students and fostering cognitive growth (Jardinez & Natividad, 2024). Despite the benefits, inclusive education also presents challenges, one of which is including gifted students. Some argue that gifted students may benefit more from specialized education rather than spending all their time in inclusive settings (Van der Meulen et al., 2014). Before proceeding to examine the placement of gifted students, it is important to look at a definition of giftedness.

Defining gifted students is complex, as they exhibit a wide range of characteristics, primarily related to intellect and learning behavior (Carman, 2013; Tirri & Lane, 2017; van Gerven, 2021; Quintero-Gaméz & Sanabria-Z, 2024). The discussion about the definition of gifted students results in a lack of standardization in tests (Quintero-Gámez & Jorge Sanabria-Z, 2024). An example of this lack of

testing standardization is the nine different identification methods described by Carmen (2013). In practice, intelligence tests are most commonly used (Quintero-Gámez & Jorge Sanabria-Z, 2024). However, these intelligence tests are limited by the fact that they do not reflect the wide range of characteristics. (Callahan et al., 2017).

As previously stated, the inclusion of gifted students has its challenges. The inclusion of gifted students is influenced by numerous misconceptions (Tirri & Laine, 2017), one of which is a hostile attitude towards gifted students and specific education for them (Persson, 2010; Van der Meulen, 2014). Teachers may have an ambivalent view of gifted students, resulting in insecurities for both teachers and students (Reis-Jorge et al., 2021). These misconceptions can result in teachers spending less time adjusting for gifted students than for low-achieving students (Tirri & Laine, 2017; Van der Meulen et al., 2014). However, gifted students have the same intensity of educational needs as students traditionally seen as students with special needs (van Gerven, 2021). According to Van der Meulen et al. (2014), there are at least five important differences between the learning of gifted and regular students: gifted students need less time to study new material, need less repetition, perceive materials on a more abstract level, struggle to shift from topics they find interesting, and reason 'top down' instead of 'bottom up'. In brief, adjusting for gifted students is important, but not always done.

As adjusting for gifted students is not always done, this raises the question of whether a regular classroom is the best place for gifted students. Plucker and Callahan (2014) describe that extensive research indicates that a regular classroom, without interventions and differentiation for gifted students, provides little challenge to gifted students. This occurs when policies are based on the assumption that regular differentiation, without additional modifications, is enough for gifted students. In this context, a carefully designed curriculum delivered in a separate classroom results in greater learning gains for gifted students (Plucker & Callahan, 2014). Similar results are found by Garcia-Perales and Almeida (2019), who illustrate that, in an education system that insufficiently supports gifted students, an

enrichment program leads to a more effective addressing of the special needs of gifted students. In summary, even though inclusion is an important international goal, inclusive education might not always be the best fit for gifted students.

There are several approaches to gifted education, ranging from full inclusion to rejection of inclusion (Jardinez & Natividad, 2024). One of these approaches is the in-and-out inclusive education model, where students benefit from both special and regular education settings (Jardinez & Natividad, 2024). This is often described as either pullout programs or enrichment programs. A pullout program is defined as an arrangement where gifted students get most of their instruction in a heterogeneous classroom, and periodically get taught in a separate setting with gifted peers. (Vaughn, 1991; Delcourt et al., 2007).

Countries approach gifted education in various ways. In the Netherlands, for instance, schools have the autonomy to decide how they organize education for gifted students, for example by creating special classes or modifying classroom instruction (Doolaard &Oudbier, 2010; Ministerie van Onderwijs, Cultuur en Wetenschap, 2021). Although segregated education exists (van Gerven, 2021), many gifted students are included in pull-out programs. Similarly, in the United States, the No Child Left Behind Act of 2002 states that gifted students need “services or activities not ordinarily provided by the school” to develop their high achievement capabilities (Callahan et al., 2017, p. 21). A survey by Callahan et al. (2017) found that half of the schools use pullout classes as a dominant service delivery model. Only 5.7% provided in-class services, without clustering gifted students. Thus, pullout programs are often used within different countries to provide education to gifted students.

As previously discussed, the use of an enrichment or pullout program might lead to a more effective addressing of the specific needs of gifted students (Garcia-Perales & Almeida, 2019). Furthermore, participating in a pullout program has a positive effect on achievement (Hornstra et al., 2017) and might lead to better academic performance compared to peers (Vaughn, 1991; Delcourt et al.,

2007). Last, students in a pullout program scored better on mental wellbeing compared to students in a fully separate classroom, and had a slightly higher mathematical self-concept than students fully in regular classes (Cash & Lin, 2021).

Pullout programs appear to have positive effects on gifted students. However, many of the programs used are not empirically tested, which makes the actual effects seen in practice unclear (Callahan et al., 2015). There is a lot of ambiguity about effective practices, as gifted students form a small group, therefore, studies often have small sample sizes, and considering that there is a lot of inconsistency in the definition of giftedness (Plucker & Callahan, 2014). While gifted programming frameworks exist, like the Enrichment Triad Model and the differentiation model, not all schools use these frameworks to guide their gifted programs (Doolaard & Oudbier, 2010; Callahan, 2017). In a further evaluation of gifted programs in the United States, Callahan (2017) found that one-quarter of the respondents did not have specific curricular materials in their gifted programs, and fewer than 50% of the districts require a gifted program evaluation. This suggests that many programs are neither tested nor evaluated, and that tested programs are not always used.

The last review on pullout programs was performed by Vaughn (1991), who found small to medium effects on the academic performance of gifted students. However, gifted education has changed significantly since that time. Since 2008, gifted education has emerged as an important political topic, due to an economic regression showing the importance of talent development (De Boer et al., 2013; Avcu & Er, 2017). At the same time, educational policy aimed at inclusion has developed, which means regular education might be more suitable than it was in 1991. A more recent review was written by Kim (2016), but this did not differentiate between different types of enrichment programs, such as those occurring during and outside of school time. Thus, it is relevant to perform a new review on the effectiveness of pullout programs.

To contribute to the debate of the best placement for gifted students, this systematic literature review will focus on the effects of pullout classes as part of mainstream education, by answering the following research question: *What is the effect of pullout classes as part of mainstream education on the academic performance of gifted students in primary school?* In the articles included in this review, giftedness has to be clearly defined, grounded in a conceptual model. Choosing one model or criterion leads to the exclusion of a lot of valuable research. The focus will be on pullout programs during school time, where students get pulled out for a certain amount of time each week. It is important to note that a pullout program is a form of a gifted program and a form of an enrichment program, but it is not the same as the term enrichment program also includes after-school and summer programs.

This review will contribute to the knowledge about pullout programs, specifically focusing on academic outcomes and student performance in schools. Furthermore, possible factors contributing to the effects will be analyzed in this review. This review will examine a way of meeting the needs of gifted students that is not full inclusion, and thus partly help answer the bigger question: Is full inclusion really what is best for gifted students at this time?

Method

The research question is addressed through a systematic literature review. This method is a reproducible approach to answering a research question, with the primary objectives of reducing bias risk and increasing transparency (Lame, 2019). By evaluating existing studies and synthesizing their findings, research gaps can be identified to show what more needs to be known (Gough et al., 2017). In the context of this research, synthesizing knowledge regarding the effectiveness of pullout programs facilitates a better comparison with other forms of gifted education. The procedure of this review adheres to the PRISMA protocol, a reporting guideline for systematic reviews, created to improve transparency and quality (Page et al., 2021). The topic was defined using the PICO framework (Hosseini et al., 2024):

- Population: gifted students
- Intervention: Pullout program
- Comparison: full inclusion
- Outcome: academic results

Data search

To determine which articles to include in this review, several inclusion and exclusion criteria were established. These criteria were structured according to the criteria of Lindner et al. (2023);

- Relevance to the topic: The article focuses on pullout programs for gifted students. Programs only used during the COVID-19 period will be excluded.
- Date of publication: The article was published between 2010 and February 2025.
- Publishing medium: The article was published in a peer-reviewed journal. It is accessible through used search engines.
- Language: The article is written in English or Dutch.
- Design: case studies are excluded. Academic performance is assessed with a test using standard measures, for example, regular state tests.
- Setting: The focus of this study is primary education. Pullout programs were organized during school time, for a certain amount of time each week. Summer programs, after-school programs, and temporary courses are excluded, as are programs targeting the regular classroom.
- Sample: The sample of the study includes or consists fully of gifted students. Giftedness is clearly defined in the article, based on a model that includes intelligence as one of the criteria.

To identify the relevant articles, a search was performed in the ERIC and PsycINFO databases. After the first relevant results were found, these were entered into the database “Web of Science” to perform a forward search. The references of screened articles were also screened to identify other relevant results, thus conducting a backward search. The articles found with forward- and backward searches had to

meet the same inclusion and exclusion criteria. The search items initially used in ERIC and PsycInfo were as follows:

- Gifted or gifted children or gifted youth or gifted and talented or gifted student* or giftedness
- Enrichment activities or gifted education programs or gifted education or pull-out program* or pull out program* or in-and-out
- Academic achievement or academic performance or academic success or grades or GPA or academic outcome* or achievement
- Primary school or elementary school or primary education or elementary education

The search was conducted using the PRISMA method (Page, 2020), which began with screening titles, followed by screening abstracts of relevant articles to determine which articles should be included. The selected studies were then subjected to further analysis, which consisted of four steps (Lindner et al., 2023): first, identifying the theoretical construct, second, describing the characteristics of the study, third, compiling similarities and differences between studies, and as a final step, writing the data synthesis. The first three steps, analyzing the data, were done using a table in Excel. Additionally, Zotero was used to annotate the articles.

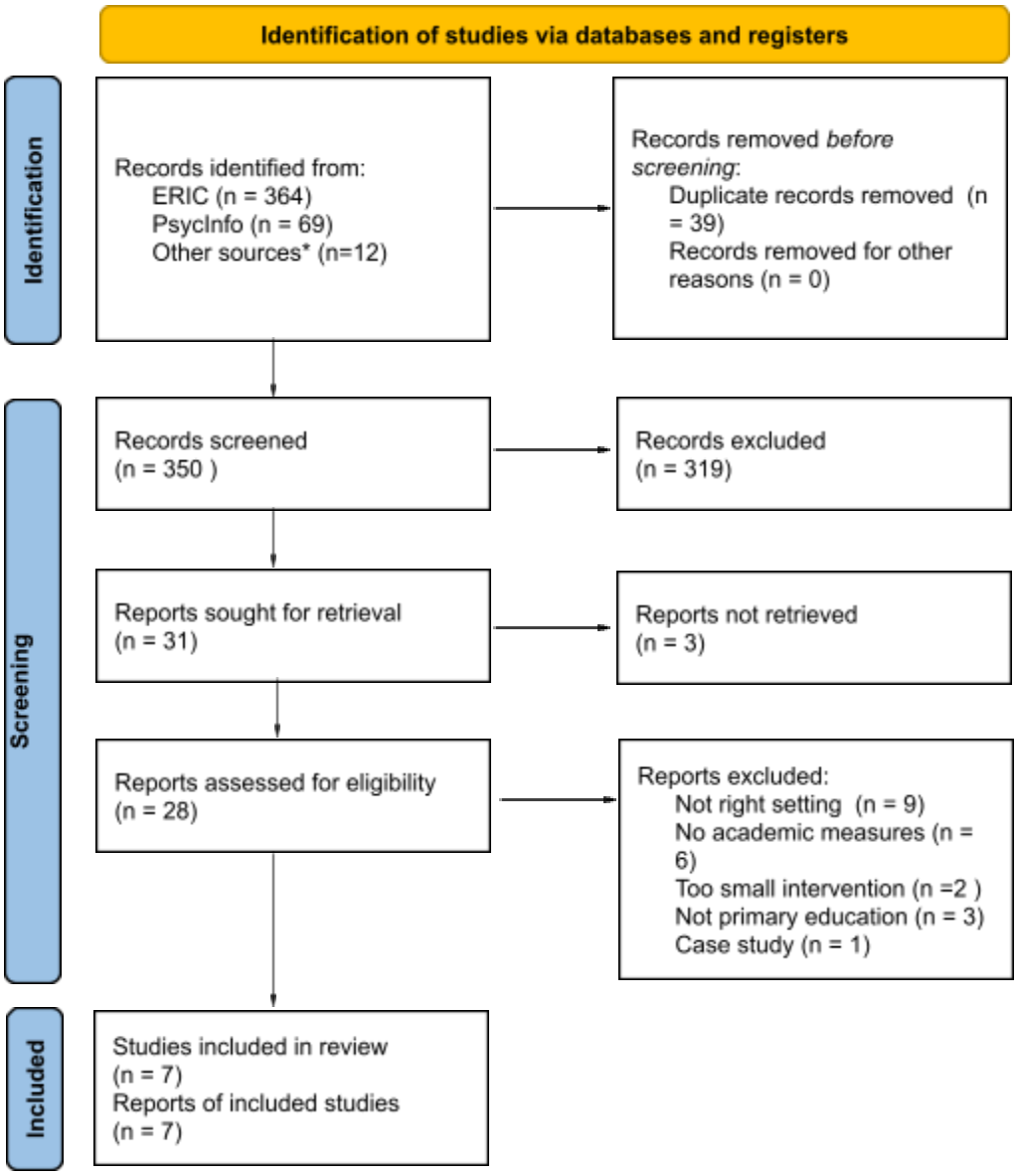
To analyze the data, a qualitative synthesis will be conducted by grouping the information from the articles according to the type of program and the outcome measures used. For this, a synthesis matrix will be used. After that, a thematic analysis is performed to identify themes within these studies that are relevant to the research question.

Data extraction

The input of the search string in ERIC led to 364 results, and in PsychINFO to 69 results. 37 duplicate articles were removed before screening. Forward and backward reference searches led to 12 additional articles, of which one made the final selection. In total, 31 articles seemed relevant based on the abstract and title, and were sought for retrieval. However, three were not available.

Figure 1

Prisma 2020 Flow diagram



** articles found through forward and backward reference searching*

Ultimately, 28 articles were screened. Of these 28 articles, nine were excluded because they did not focus on gifted students in pullout programs, but rather on programs for gifted students either in regular or completely separate classrooms. Six articles were excluded because they did not use standard testing measures, but rather a self-made test or non-academic measures. Two articles were excluded

because they were intervention studies, which only lasted a short time. These did not meet the criteria as a continuous program for gifted students. Three articles focused on middle- or secondary education and were thus excluded. Lastly, one article was excluded because it was a case study. In the end, seven articles met the inclusion criteria. This selection process is portrayed in Figure 1.

Among these seven articles is the study by the NCRGE et al. (2019). This article was a Brief on the results of a full cohort of gifted students in three different states, rather than focusing on a specific program. Further publications derived from this data, also published by the NCRGE, were included in the results. These were found on the NCRGE website and mainly consisted of PowerPoint presentations on this data.

Quality assessment

While there is a lot of research being done, not all sources are trustworthy, valid, and reliable (Hong et al., 2018). Therefore, it is important to check the quality of the studies included in this review. In order to nuance the findings of the included studies, quality assessment was thus done. As the search did not lead to many results, articles were not included if they did not meet the quality criteria.

The quality assessment was performed using the mixed methods appraisal tool, or MMAT (Hong et al. 2018). The article by the NCRGE et al. (2019) was not assessed, as it did not fit the initial screening criteria, consisting of clear research questions and matching collected data. Three other articles did not clearly state their research question, but the context made clear that they were looking into the effects of their program.

The MMAT has separate criteria for different types of research. Two articles were screened using the criteria of Randomized controlled trials (Callahan et al., 2015; Gubbels et al., 2014). Neither met the criterion of blinded outcome assessors. In addition, the study of Gubbels et al. (2014) did not meet the criterion of comparable groups at baseline, which they corrected for by including age as a covariate in the rest of the study. The remaining four articles were screened using the criteria for non-randomized

studies. The article by Garcia-Perales and Almeida (2019) stood out as there was no clear description of confounder control. Caution should be used when interpreting the results of this article. In addition, not all articles had clear guidelines for their programs. This made it difficult to assess the criterion of program fidelity. However, as only one program followed strict program guidelines and reported adherence (Callahan et al., 2015), the other studies had less strict guidelines or developed the program during implementation, which makes program fidelity a less relevant criterion.

Results

The results section begins with the descriptive statistics of the included studies. Following this, the thematic analysis is presented. This will start with the model of giftedness and the knowledge gap used among the different studies. Then, academic achievement and its subthemes are described. Next, other outcomes of the programs are addressed. Finally, potential factors that may influence the outcomes of the pullout programs are identified.

Descriptive statistics

Seven studies met the inclusion and exclusion criteria, as shown in Table 1. These seven studies were conducted in five countries, including the United States (n=2), the Netherlands (n=2), Spain (n=1), Germany (n=1), and China, specifically Hong Kong (n=1). Out of these studies, six studies used pre- and posttests as measurements. The study by the NCRGE et al. (2019) analyzed longitudinal data of a cohort for three years, therefore, they did not use a specific pre- and posttest. Of the six studies using pre- and posttests, five used a control group for comparison. Only Van der Meulen et al. (2014) did not use a control group, as they followed an exploratory, uncontrolled open trial design.

All studies that included control groups included students from different schools in the control group (n=5). In one case, there was a distinction between two control groups: one from the same school as the experimental group, and the other from other schools. The students from other schools were

Table 1

Overview of included studies

Study	*Study design*	*Selection and design Pullout program*	*Academic measures*	*Findings*
Callahan et al. (2015) United States	Cluster-randomized experiment with 2905 students in grade 3	- Two specific units - 45-60 minutes, twice a week - Selection by schools procedure	ITBS survey battery reading subtest and standards-referenced posttest	EG scored significantly better at academic outcome measures
Garcia-Perales and Almeida (2019) Spain	Quasi-experimental: pretest-posttest with control group: 9 EG, 27+9* CG, in grades 2-6	- Horizontal enrichment - 3 weekly sessions of 2 hours, of which 2 outside the classroom - Diagnosed by school counseling services	Class records, school performance, IQ scores	- 4 EG improved school performance - 3 CG (of 27) improved performance
Golle et al. (2018) Germany	Quasi-experimental: pretest-posttest with matched control group with 423 EG, 2328 CG, in grade 3	- HCAP-program - Around 2 hours each week - Teacher nominated	Intelligence, school grades in mathematics, and German	- EG significantly better German and mathematics grades
Gubbels et al. (2014) Netherlands	Experimental: pretest-posttest with control group with 32 EG, 34 CG, in grades 5-6	- Triarchic teaching - 3 hours each week - Teacher nomination based on above-average intelligence, motivation, and absence of clinical diagnoses	Intelligence, triarchic abilities (translation of the American Aurora Assessment Battery)	- EG: Increase in practical intelligence - Analytical intelligence stable (attributed to ceiling effect)

Study	*Study design*	*Selection and design pullout program*	*Academic measures*	*Findings*
NCRGE et al. (2019) United States	Analysis of survey and longitudinal data with 1 cohort in 3 states (362,254 students total), followed grades 3-5	- Different programs for every school - Selection based on individual school's procedure	Longitudinal student-level academic achievement data	- No effect of language arts and mathematics gifted classes on the academic achievement - Gifted students start 2 years ahead in 3rd grade, but grow more slowly than regular students in grades 3-5
Shek et al. (2022) China	Quasi-experimental: pretest-posttest with 2551 in total sample, 176 in pullout, Grade 3-6 primary school, and 1-3 secondary school	- Enrichment and acceleration - Selection based on school performance, parent-teacher- or self-nomination	Multiple intelligences, academic performance	- Pullout program low-to-medium effect sizes on academic performance, logical–mathematical intelligence, visual-spatial intelligence, and intrapersonal intelligence
Van der Meulen et al. (2014) Netherlands	Exploratory: uncontrolled open trial design with 89 EG, grades 3-5	- Implementation of British "DWS" program - 1 day a week - Extensive selection procedure*	Academic results in language, reading, and mathematics	- At-risk group had large, significant increase language results, large, not significant effect on reading, and no significant effect on mathematics

EG stands for experimental group, and CG stands for control group

**Two control groups were used: 27 students in the same school, 9 students in other schools.*

***Full procedure not described, see van der Meulen et al. (2014).*

matched based on the selection criteria of the study (Garcia-Perales and Almeida, 2019). The students of the experimental group were from the same school in the case of Carcia-Perales and Almeida (2019), but from different schools in a region in all other studies (n=4). This demonstrates how representative the sample is.

Most studies focused on primary education only (n=6), only Shek et al. (2022) focused on primary as well as secondary education. The time spent in pullout programs was reported by five out of seven studies, varying from two hours to a full school day (6 hours and 45 minutes). In one study, this depended on the specific enrichment programs followed. Two studies did not report the amount of time spent in the pullout program (NCRGE et al., 2019; Shek et al., 2022).

In the intervention studies, there is a big variety in the type of gifted program being examined. One study focused on modules, based on the CLEAR model (Callahan et al., 2015). Additionally, two articles build lessons based on theory (Gubbels et al., 2014; Van der Meulen et al., 2014). In contrast, two articles included many different enrichment activities, without clearly stating what specific theory was behind this (Garcia-Perales and Almeida, 2019; Shek et al., 2022). Lastly, Golle et al. (2018) studied a grassroots program, which meant that there was a lot of variety in the enrichment activities the students received. Consequently, even within pullout programs conducted during school hours, there are a lot of differences, both in foundation and implementation.

Model of giftedness

As discussed in the introduction, there is a lot of variation in the definition of giftedness. This is visible in the studies included in this review, as none of them employ the same definitions or selection criteria. Two articles choose a specific model: Gubbels et al. (2014) use the triarchic theory of intelligence, and Shek et al. (2022) the three-ring conception of giftedness. Both have a more extensive selection procedure than teacher nomination, although this is one part of it.

The remaining articles describe either the disarray in definitions (Golle et al., 2018), the different

characteristics associated with giftedness (Van der Meulen et al., 2014), the possible emotional and behavioral difficulties (Garcia-Perales and Almeida, 2019), or the various challenges of gifted education (Callahan et al., 2015). In the selection process for the pullout program, either teacher nomination is used (Golle et al., 2018), selection is done by school psychologists (Garcia-Perales and Almeida, 2019), or it depends on the state (Callahan et al., 2015; NCRGE et al., 2019). This variety in view of giftedness and way of selection suggests that there might be students selected by some articles that would not have been selected by others. The samples of different studies might thus show differences that hinder direct comparability.

Addressed knowledge gap

In addition to the variety in definition and selection of gifted students, another challenge in comparing pullout programs is the lack of empirical knowledge about, and wide variety of, pullout programs. While all included studies aim to address this gap, their method varies. The study of Golle et al. (2018) notes that statewide programs are barely studied, and thus reviews a program as currently used, which has a lot of variety in terms of implementation. This is comparable to the analysis of longitudinal data conducted by the NCRGE et al. (2019). In contrast, the other studies created and implemented new pullout programs (Garcia-Perales and Almeida, 2019; Gubbels et al., 2014; Shek et al., 2022). Not all studies clearly described the characteristics and structure of their respective program. Lastly, Callahan et al. (2015) did not create a new pullout program, but created a curriculum to be implemented in existing groups. This variety in form and content of the program makes it difficult to determine what led to the results of the program, and thus, which part of the program was effective.

Academic results

Having discussed the variety in pullout programs included in this study, this next section will describe and compare the academic outcomes of the different programs. An important note is that not

all articles focused on the same type of academic outcomes, for example, school subjects, and even when focusing on the same type of outcome, they used different measurements.

General academic achievement

When reporting the effects of a program, not all articles specify how they determine academic improvement. Two articles used a general measure for academic achievement, both reporting positive effects. In the study of Garcia-Perales and Almeida (2019), school results were gathered through class records. They found that 44.4% of students in the experimental group improved their school results, compared to only 8.3% in the control group. Additionally, in the control group, one student regressed. A similar view of academic performance is presented by Shek et al. (2022), where the experimental group had a significant improvement in academic performance ($p<0.000$), when compared to a control group. However, it is important to note that the control group did not match the demographics of the experimental group. Last, the study of Callahan et al. (2015) looked partly at research skills and saw a significant difference after participating in their program, looking at a standards-based test.

Mathematic achievement

Three studies examined the effects of pullout programs on mathematics scores, yielding mixed results. Two articles looked at specific effects of a certain program on math performance. Among these two, only one found a statistically significant effect on mathematics grades (Golle et al., 2018). Interestingly, the article that found significant improvement in mathematics did not see improvement in non-academic areas, whereas the article that did not find a significant effect in mathematics did find significant effects in many other areas, such as academic self-concept (Van der Meulen et al., 2014). These non-academic measures and possible factors contributing to these effects will be discussed later. The study by the NCRGE et al. (2019) found that pull-out programs in the three states in their research had less focus on mathematics, more on creativity and critical thinking skills. While the gifted students in

this study are far ahead in grade three, they grow more slowly than regular students in grades three to five.

Achievement in language arts

Three programs reported on language arts, all showed a significant increase in scores on either the total experimental group or the at-risk group. The program of Callahan et al. (2015) focused partly on poetry and built a standards-referenced test measuring the improvement of the students. They found that students who participated in their course scored significantly higher on the test. The combination of this result and the result on the research unit led to their conclusion that the experimental group scored significantly better on academic outcome measures. In addition, two other studies examined the effects of pullout programs on grades in students' native languages. Golle et al. (2018) found significantly improved German grades in their experimental group, and Van der Meulen et al. (2014) found a large, significant increase in the Dutch language results of their at-risk group. However, this increase was not observed in their total group, possibly due to ceiling effects.

Reading improvement

Only one study specifically examined the reading skills of students in their program (Van der Meulen et al., 2014). Although they did not find a significant effect of the program in both the total experimental group and the at-risk students, they did note that there was a large difference in scores before and after the program for the at-risk group. This was based on school grades, reported by the students' teachers.

Intelligence

Three articles reported the impact of their programs on the intelligence of their students. All found mixed results, as some subtests improved, but not all. The tests used different subscales and are thus not easily comparable. Golle et al. (2018) reported no differences in either fluid or crystallized intelligence after participants followed the program, thus finding no significant effect on intelligence.

Gubbels et al. (2014) reported a significant change in practical, but not analytical or creative intelligence. Finally, Shek et al. (2022) used a Student multiple intelligence profile and reported a higher level of logical-mathematical intelligence, visual-spatial intelligence, and intrapersonal intelligence, but noted that verbal-linguistic, musical, bodily-kinesthetic, and naturalistic intelligences did not show similar improvements.

Other outcomes from programs

This review focuses on academic improvement achieved through pullout programs. However, it is relevant to consider this improvement within its broader context. Five of the studies included in this review reported on non-academic measures, with four of them finding significant positive results. These articles either had improved behavioural and social-emotional skills as a primary goal (Garcia-Perales and Almeida, 2019; Van der Meulen et al., 2014) or reported on them in addition to academic measures (Golle et al., 2018; Gubbels et al., 2014; Shek et al., 2022). Non-academic measures, such as motivation, may also contribute to academic improvements (Gubbels et al., 2014). A clear example of not having academic results as the primary focus is shown in the conclusion of Garcia-Perales and Almeida (2019, p. 44):

> In this study, we have seen that the enrichment program for highly able students during school hours helped them improve their adaptation in general and on a personal, school, and societal level, with some of the students even improving their school performance.

The found non-academic results are categorized and explained further:

Self-concept

Three articles examined self-concept, yielding mixed results. Two studies found significant enhancements in the experimental group (Gubbels et al., 2014; Van der Meulen et al., 2014), while one study reported no significant effect (Golle et al., 2018). Additionally, Shek et al. (2022) observed improvements in self-efficacy and self-acceptance.

Attitude towards school and learning

Three articles reported on attitudes towards school and learning. While none found significant improvement in the total experimental group (Golle et al., 2018; Gubbels et al., 2014), Gubbels et al. (2014) found that motivation for school and attitude towards science decreased in the control group, but remained stable in the experimental group. This suggests that the program helped maintain motivation. Furthermore, Van der Meulen et al. (2014) reported a medium positive effect on school enjoyment for the at-risk students in the experimental group.

Effects on student creativity

Two articles reported on student creativity, presenting contrasting results. While Shek et al. (2022) found a higher level of creativity in their program, Golle et al. (2018) found no significant results.

Effects on behaviour in school

Two articles reported on behaviour in school, finding positive results on self-reported, but not on teacher-reported tests. In contrast to the results on behaviour in school, Golle et al. (2018) reported no significant effect on self-reported self-control.

Effects on social-emotional skills and well-being

The findings on well-being and social competence vary. While Gubbels et al. (2014) found no effect on well-being, Shek et al. (2022) observed a higher level of psychological well-being, personal growth, purpose in life, self-acceptance, and satisfaction with life in students participating in their program. Garcia-Perales and Almeida (2019) found improved personal adaptation. Regarding social competence, only Garcia-Perales and Almeida (2019) found a significant effect, while Golle et al. (2018) and Shek et al. (2022) did not.

Factors of improvement

Thus far, the various outcomes of the included programs have been discussed, showing that all included intervention studies showed improvement in either some or all measured areas, both

academically and non-academically. While it is important to study the effects of the program, it is equally important to analyze the factors contributing to these effects. What follows is a brief overview of the different potential factors that are suggested as working mechanisms in the intervention. These factors can be embodied in a rich curriculum, individual attention, motivation, and family support.

The first factor is that a rich curriculum leads to better academic achievement. An example of this is the CLEAR model, used by Callahan et al. (2015) to develop their program. In this model, content standards are adhered to "at a level of content and skill differentiation beyond what other students could, should, or would engage in" (Callahan et al., 2015, p. 159). Similar arguments are given by Gubbels, as their program exploited analytical, practical, and creative abilities (Gubbels et al., 2014), and Shek et al. (2022), who describe a focus on higher cognitive abilities.

Subsequently, as this curriculum is given in a different setting, the second factor is that more individualized guidance in the program leads to the improvement of the students. The responsive instruction with distinct educational attention, as well as the ability to move more at the students' own pace, might be one of the reasons gifted students benefit from these pullout programs (Callahan et al., 2015; Garcia-Perales and Almeida, 2019; Shek et al., 2022).

A third factor is that participation in a pullout program leads to more motivation, which indirectly leads to higher scholastic achievement. This is described by Gubbels et al. (2014, p. 392): "positive effects on motivation and self-concept were found. Motivation plays a crucial role in the process through which one reaches excellence, and a positive self-concept is reciprocally related to high achievements." One factor in increasing motivation might be the mix with similar students (Gubbels et al., 2014). Next to this, the identification-commitment model is proposed by Golle et al. (2018), which describes that an increased identification with school, which is achieved by spending time on school-related subjects, predicts positive academic as well as non-academic outcomes. Interestingly enough, the program of Golle et al. (2018) did not lead to these positive non-academic outcomes.

A last factor for improvement is a possible change in the students' environment. Families might become either more supportive or achievement-oriented when their child is nominated for a gifted program (Golle et al., 2018).

Discussion

This review tried to answer the question: "*What is the effect of pullout classes as part of mainstream education on the academic results of gifted students in primary school?"* Participation in pullout programs demonstrates a positive impact on academic performance. Furthermore, non-academic factors improved, albeit with more mixed results. Possible factors for improvement might be the use of a rich curriculum, individualized attention, increased motivation, and changed family support.

A previous meta-analysis in this field showed small to medium effects on academic achievement of pullout programs (Vaughn, 1991). In this review, similar positive effects are shown. In addition, while Vaughn found no positive effects on gifted students' self-concepts, this review showed that three out of four studies looking into this area did find significant effects. However, as this review did not have the capacity to perform a meta-analysis, this was not performed. A meta-analysis would have made a quantitative analysis possible, while this study only included a qualitative analysis (Lame, 2019). A quantitative analysis would make the comparison to Vaughn's findings easier.

Additionally, a more recent review performed by Kim et al. (2016) reported positive effects of various types of enrichment programs on gifted students, but it did not distinguish among these different types. This study specifically examined pullout programs, a type of enrichment program, showing that this specific form has a positive effect on gifted students.

Several aspects could be responsible for the positive results in the reviewed studies. One important factor is good communication between teachers in pullout and regular classrooms (Van der Meulen et al., 2014). Next, as described in the results section, a rich curriculum leads to better results.

This reflects the concept that in gifted education, the focus is often on enrichment, which means introducing advanced content and broader topics to develop higher-order thinking skills and provide challenges (Kalaji & Alborno, 2023). Finally, in the results, motivation was identified as an important factor. One key factor influencing motivation is need satisfaction, which Hornstra et al. (2022) found to be greater for gifted students in pullout classes compared to regular classes.

An interesting finding in this study is that, although all included intervention studies showed some improvement in their experimental group, often significant effects were not found on all measures. The grass-roots program by Golle et al. (2018) stands out specifically, as they found no significant effect on any of the non-academic measures, contrasting all other articles including such measures. A closer examination of their program shows that it had weak theoretical links between the program goals and specific outcomes, as a lot of variety existed in the content of the programs of the participating students. They describe: “local units have different teachers/instructors and different student compositions, might emphasize different aims of the HCAP, and might be organized differently” (Golle et al., 2018, p. 394). This seems similar to the results in the study of the NCRGE et al. (2019), which showed that gifted students start 3 years ahead of their peers in third grade, but grow more slowly in the following two years. In the gifted programs they analyzed, there was often also no clear guidance in the gifted program. A possible explanation is the relationship between program fidelity and student learning: “ Low fidelity increases the likelihood of weak student outcomes, and fidelity itself is predicted by both study design characteristics and program features”(Hill & Erickson, 2019, p. 596). This relationship between fidelity and program outcomes is also found in the research of Callahan et al. (2015), which was included in this review.

A reason for the lack of improvement across all academic measures might be that, when standards-referenced tests are used, there is the possibility of a ceiling effect. This means that the test could not measure growth, as there is not enough room for improvement (Matthes et al., 2024). This

was a possible reason for the lack of improvement of the total group in the study of Van der Meulen et al. (2014). In this study, they distinguished the at-risk group based on problematic scores at pretest. While they found an improvement in the academic achievement of the at-risk group, specifically a significant, large increase in language results and a non-significant, large effect in reading, they did not find an improvement in the total group. Not all articles report on this possible ceiling effect, however, this could be a possible factor in other studies using these kinds of tests (Matthes et al., 2024).

Although inclusion in education is widely discussed in international policy and seen as an important goal (UNESCO, 1994; United Nations, 2006; United Nations, 2020), the current review shows that, often in gifted education, the focus is not on inclusion. There is a notable contradiction: while there is a growing emphasis on inclusion, simultaneously, more segregated education for gifted students is created (van Gerven, 2021). Of all articles included in this review, only Van der Meulen et al. (2014) and Gubbels et al. (2014) dive into inclusive education as an option. Both find inclusion insufficient, as there is a mismatch between cognitive abilities and the instructional environment. This is due to a lack of teacher training and a focus on weaker learners. This mismatch has several negative consequences, such as a lack of demanding learning experiences, boredom, depression, and a lack of confidence in challenging tasks, which then leads to underachievement.

While the other articles in this review do not specifically mention inclusion as an option, Garcia-Perales and Almeida (2019) do specifically describe limitations of implementing curricular modifications in the regular classroom, for example, lack of teacher training and prejudice towards gifted students. The negative effects described in these articles are also found in other studies. More specifically, the notion that inclusion is lacking is further supported by Yildiz (2022), the lack of teacher training by De Boer et al. (2013) and Tirri and Laine (2017), and Persson (2010) describes the hostility and prejudice towards gifted students.

There are several options for gifted education. Alongside full inclusion, there are also fully separate classrooms. Van der Meulen et al. (2014, p.296) describe that both approaches have negative side effects, which leads to the idea that the pullout program can counter these negative effects:

> Solely adaptations within the classroom seem to have little academic effects. On the other hand, special schools for gifted children, giving the most academic effects, can cause lower self-concepts and less preference for challenging tasks. These findings give support for a form that is in between these extremes

This idea is supported by the research of Hornstra et al. (2017), which indicates that students in a part-time program had the most favorable emotions, compared to both full–time programs and regular education. However, these emotions might not spill over to the regular school.

In the quote above, Van der Meulen et al. (2014) describe that solely adaptations in the classroom do not have enough effect on gifted students. In the search for gifted pullout programs, however, many interventions for whole classrooms arose. A relevant question to ask is thus whether the effects of pullout programs still seem better than full classroom interventions. A review by Marsili et al. (2023) revealed that these types of interventions led to mixed results on learning outcomes. Nonetheless, they identified differentiation by teachers as a promising approach. In addition, a curriculum specifically designed for gifted students might be beneficial even if implemented in the regular classroom. Callahan et al. (2015), whose curriculum was evaluated in the results, described this as a possible use for the education now designed for gifted programs.

Conclusion

This systematic literature review analyzed pullout programs for gifted children to see if the academic scores of gifted children improved. It was limited, as the primary focus was on academic performance. Programs only looking at social-emotional measures were thus excluded. A weakness of this study was that not all studies had the same selection criteria for gifted children, which means that

not all children would have matched the gifted models, for example, because students were selected by their teacher (Golle et al., 2018). Lastly, studies with negative results are less published, which might have led to a bias in the results (Lame, 2019).

This review has shown several opportunities for follow-up research. First, a meta-analysis of the pullout programs, including programs with social-emotional measures, is needed, as the last meta-analysis in this field was published by Vaughn (1991). This meta-analysis would allow a quantitative analysis of the data, which could combine the individual studies using the common metric effect size (Therrien et al., 2010). While this review has clearly shown positive results, the statistics in a meta-analysis might lead to a better view, especially comparing pullout programs to other forms of gifted education.

Furthermore, looking at the specific pullout programs, barely any broad programs are assessed, mostly small interventions, while the broad programs are used in education (Golle et al., 2018). Programs that have shown positive results are rarely implemented in practice; rather, practices are based on individuals' opinions (Siegle et al., 2024). This gap is also identified by the NCRGE et al. (2019), who describe that teachers have a lot of freedom in choosing the content of their gifted program. Moving forward, there is a need for the assessment of the broad programs currently used (Golle et al., 2018) and the implementation of research-based programs (Siegle et al., 2024). Lastly, none of the included studies performed follow-up tests. This could be beneficial to connect the short-run outcomes to long-term effects, on measures such as academic outcomes, intelligence, and motivation (Watts et al., 2019).

This study showed that, even though international policy aims for inclusive education, gifted students often benefit from participating in pullout programs, both academically and on non-academic measures. Thus, participating in a pullout program also benefits the students when they are in their regular classroom.

References:

Avcu, Y. E., Er, K. O. (2017). Education of gifted and talented individuals in Germany, the Netherlands, Sweden and Turkey: An investigation of education policies and its implementations. *International Online Journal of Educational Sciences*, 9(4), 1154–1170. https://doi.org/10.15345/iojes.2017.04.018

Brown v. Board of Education, 347 U.S. 483 (1954).

Callahan, C. M., Moon, T. R., Oh, S., Azano, A. P., & Hailey, E. P. (2015). What works in gifted education: Documenting the effects of an integrated curricular/instructional model for gifted students. *American Educational Research Journal, 52(1)*, 137–167. https://doi.org/10.3102/0002831214549448

Callahan, C. M., Moon, T. R., & Oh, S. (2017). Describing the Status of Programs for the Gifted: A Call for Action. *Journal for the Education of the Gifted, 40(1)*, 20-49. https://doi.org/10.1177/0162353216686215

Carman, C. A. (2013). Comparing apples and oranges: Fifteen years of definitions of giftedness in research. *Journal of Advanced Academics, 24(1)*, 52-70. https://doi.org/10.1177/1932202X1247260

Cash, T. N., Lin, T.-J., National Association for Gifted Children, & The Council of State Directors of Programs for the Gifted. (2021). Psychological well-being of intellectually and academically gifted students in self-contained and pull-out gifted programs. *Gifted Child Quarterly, 66(3)*, 188–207. https://doi.org/10.1177/00169862211032987

De Boer, G. C., Minnaert, A. E., & Kamphof, G. (2013). Gifted education in the Netherlands. *Journal for the Education of the Gifted*, *36(1)*, 133–150. https://doi.org/10.1177/0162353212471622

Delcourt, M. A. B., Cornell, D. G., & Goldberg, M. D. (2007). Cognitive and affective learning outcomes of gifted elementary school students. *Gifted Child Quarterly*, *51(4)*, 359–381.

https://doi.org/10.1177/0016986207306320

Doolaard, S., & Oudbier, M. (2010). *Onderwijsaanbod aan (hoog)begaafde leerlingen in het basisonderwijs.* Inspectie van het onderwijs. https://research.rug.nl/nl/publications/c1a60c39-e3be-4ff7-b037-489eba9a7fd3

García-Perales, R., & Almeida, L.-S. (2019). An enrichment program for students with high intellectual ability: positive effects on school adaptation. *Comunicar: Media Education Research Journal*, *27(60)*, 39–47. https://doi.org/10.3916/C60-2019-04

Golle, J., Zettler, I., Rose N., Trautwein U., Hasselhorn, M., & Nagengast B. (2018). Effectiveness of a "Grass Roots" statewide enrichment program for gifted elementary school children, *Journal of Research on Educational Effectiveness*, *11(3)*, 375-408. https://doi.org/10.1080/19345747.2017.1402396

Gough, D., Thomas, J., & Oliver, S. (2017). *An introduction to systematic reviews.* SAGE Publications.

Gubbels, J., Segers, E., & Verhoeven, L. (2014). Cognitive, socioemotional, and attitudinal effects of a triarchic enrichment program for gifted children. *Journal for the Education of the Gifted, 37(4)*, 378–397. https://doi.org/10.1177/0162353214552565

Hill, H. C., & Erickson, A. (2019). Using implementation fidelity to aid in interpreting program impacts: A brief review. *Educational Researcher, 48(9)*, 590-598. https://doi.org/10.3102/0013189X19891436

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, *34(4)*, 285-291. https://doi.org/10.3233/EFI-180221

Hornstra, L., van der Veen, I., & Peetsma, T. (2017). Effects of full-time and part-time high-ability programs on developments in students' achievement emotions. *High Ability Studies, 28(2)*, 199–224. https://doi.org/10.1080/13598139.2017.1332575

Hornstra, L., van Weerdenburg, M., van den Brand, M., Hoogeveen, L., & Bakx, A. (2022). High-ability students' need satisfaction and motivation in pull-out and regular classes: A quantitative and qualitative comparison between settings. *Roeper Review, 44(3)*, 157–172. https://doi.org/10.1080/02783193.2022.2071367

Hosseini, M.-S., Jahanshahlou, F., Akbarzadeh, M. A., Zarei, M., & Vaez-Gharamaleki, Y. (2024). Formulating research questions for evidence-based studies. *Journal of Medicine, Surgery, and Public Health*, *2*. https://doi.org/10.1016/j.glmedi.2023.100046

Kalaji, G., & Alborno, N. (2023). The influence of gifted and talented programs on students' self-concept. *Athens Journal of Education*, *10(3)*, 507–522. https://doi.org/10.30958/aje.X-Y-Z

Kim, M. (2016). A meta-analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly*, *60(2)*, 102–116. https://doi.org/10.1177/0016986216630607

Lame, G. (2019). Systematic literature reviews: An introduction. In *Proceedings of the design society: international conference on engineering design*, *1(1)*, 1633-1642. https://doi.org/10.1017/dsi.2019.169

Lindner, K. T., Schwab, S., Emara, M., & Avramidis, E. (2023). Do teachers favor the inclusion of all students? A systematic review of primary school teachers' attitudes towards inclusive education. *European Journal of Special Needs Education*, *38(6)*, 766-787. https://doi.org/10.1080/08856257.2023.2172894

Marsili, F., Dell'Anna, S., & Pellegrini, M. (2023). Giftedness in inclusive education: a systematic review of research. *International Journal of Inclusive Education*, *29(4)*, 502–519. https://doi.org/10.1080/13603116.2023.2190330

Matthes, J., Schneider, M., & Preckel, F. (2024). The relation between prior knowledge and learning in regular and gifted classes: A multigroup latent growth curve analysis. *Journal of Educational Psychology*, *116(2)*, 278–296. https://doi.org/10.1037/edu0000848

Jardinez, M. J., & Natividad, L. R., (2024). The advantages and challenges of inclusive education: Striving

for equity in the classroom. *Shanlax International Journal of Education*, *12*(2), 57–65. http://dx.doi.org/10.34293/education.v12i2.7182

Ministerie van Onderwijs, Cultuur en Wetenschap. (2021, June 21). *Hoe krijgt mijn hoogbegaafde kind onderwijs?* Rijksoverheid.nl. Https://www.rijksoverheid.nl/onderwerpen/passend-onderwijs/vraag-en-antwoord/hoe-krijgt-mijn-hoogbegaafde-kind-onderwijs

National Center for Research on Gifted Education, University of Connecticut, Siegle, D., McCoach, D. B., Gubbins, E. J., & Callahan, C. (2019). *Results of four years' research at the National Center for Research on Gifted Education: National Center for Research on Gifted Education (NCRGE) brief on gifted education curriculum and gifted achievement growth of gifted students in three States.* Paper presented at 23rd World Biennial Conference on Gifted Children. (23rd, Nashville, KY, July 27, 2019) https://ncrge.uconn.edu/conference-papers/

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372(71). https://doi.org/10.1136/bmj.n71

Persson, R. S. (2010). Experiences of Intellectually Gifted Students in an Egalitarian and Inclusive Educational System: A Survey Study. Journal for the Education of the Gifted, 33(4), 536–569. https://doi.org/10.1177/016235321003300405

Quintero-Gámez, L., & Sanabria-Z., J. (2024). Global insights in giftedness research: Mapping current characteristics and challenges. *Journal of Social Studies Education Research*, *15(4)*, 197–222. https://www.researchgate.net/publication/384240865

Plucker, J. A., & Callahan, C. M. (2014). Research on giftedness and gifted education: Status of the field and considerations for the future. *Exceptional Children*, *80(4)*, 390-406.

https://doi.org/10.1177/0014402914527244

Reis-Jorge, J., Ferreira, M., Olcina-Sempere, G., & Marques, B. (2021). Perceptions of giftedness and classroom practice with gifted children -- An exploratory study of primary school teachers. *Qualitative Research in Education*, *10(3)*, 291–315. http://dx.doi.org/10.17583/qre.8097

Shek, D. T. L., Cheung, A. C. K., Hui, A. N. N., Leung, K. H., Cheung, R. S. H., & Tchounwou, P. B. (2022). Development and evaluation of a pioneer school-based gifted education program (Project GIFT) for primary and secondary students in Hong Kong. I*nternational Journal of Environmental Research and Public Health*, *19(8)*. https://doi.org/10.3390/ijerph19084832

Siegle, D., Hook, T. S., & Wright, K. J. (2024). Confronting the Gordian knot: Disentangling gifted education's major issues. *Gifted Child Quarterly*, *68(3)*, 175-188. https://doi-org.proxy-ub.rug.nl/10.1177/00169862241239617

Therrien, W. J., Zaman, M., & Banda, D. R. (2010). How can meta-analyses guide practice? A review of the learning disability research base. *Remedial and Special Education*, *32(3)*, 206-218. https://doi-org.proxy-ub.rug.nl/10.1177/0741932510361266

Tirri, K., & Laine, S. (2017). Ethical challenges in inclusive education for gifted students. *International Journal of Inclusive Education, 21(6)*, 1–15.https://doi.org/10.1108/S1479-363620170000009010

United Nations Convention on the Rights of Persons with Disabilities, December 13, 2006, https://www.ohchr.org/en/hrbodies/crpd/pages/conventionrightspersonswithdisabilities.aspx

United Nations. (2020, 29 juni). *Sustainable Development Goals (SDG 4) | United Nations Western Europe*. United Nations Western Europe. https://unric.org/en/sdg-4/

United Nations Educational, Scientific and Cultural Organization (1994). *The Salamanca statement and framework for action on special needs education*. Salamanca, Spain: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000098427

Van der Meulen, R. T., van der Bruggen, C. O., Spilt, J. L., Verouden, J., Berkhout, M., & Bögels, S. M. (2014). The pullout program day a week school for gifted children: Effects on social-emotional

and academic functioning. *Child & Youth Care Forum*, *43(3)*, 287–314. https://doi.org/10.1007/s10566-013-9239-5

Van Gerven, E. (2021). Educational paradigm shifts and the effects on educating gifted students in the Netherlands and Flanders. *Journal for the Education of the Gifted, 44(2)*, 171-200. https://doi-org.proxy-ub.rug.nl/10.1177/01623532211001452

Vaughn, V. L. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, *35(2)*, 92–98. https://doi-org.proxy-ub.rug.nl/10.1177/001698629103500208

Watts, T. W., Bailey, D. H., & Li, C. (2019). Aiming further: Addressing the need for high-quality longitudinal research in education. *Journal of Research on Educational Effectiveness*, *12(4)*, 648–658. https://doi-org.proxy-ub.rug.nl/10.1080/19345747.2019.1644692

Yell, M. (2021). Brown v. Board of Education and the development of special education. Intervention in school and clinic, *57(3)*, 198-200. https://doi.org/10.1177/10534512211014874

Yildiz, A. (2022). Examining the problem posing skills of gifted students in mathematics teaching. *Research in Pedagogy*, *12(1)*, 1–14. http://dx.doi.org/10.5937/IstrPed2201001Y