KU LEUVEN

FACULTEIT BEWEGINGS- EN
REVALIDATIEWETENSCHAPPEN

Relative age and gender effects on the new entrance exam for medicine and dentistry education

door Dina Dani
en Aurelia Steyaert

masterproef aangeboden, tot het behalen van de graad van Master of Medicine in de sportgeneeskunde

o.l.v.

prof. dr. em. W. Helsen, promotor

Leuven, 2023

FACULTEIT BEWEGINGS- EN
REVALIDATIEWETENSCHAPPEN

Relative age and gender effects on the new entrance exam for medicine and dentistry education

door Dina Dani
en Aurelia Steyaert

masterproef aangeboden, tot het behalen van de graad van Master of Medicine in de sportgeneeskunde

o.l.v.

prof. dr. em. W. Helsen, promotor

Leuven, 2023

Opgesteld volgens de richtlijnen van het tijdschrift *Frontiers in Psychology*

Woord vooraf

Van harte willen we prof. dr. em. W. Helsen bedanken voor zijn intensieve begeleiding en waardevolle feedback bij de realisatie van deze masterproef. Hij vervulde zijn rol duaal als zijnde de promotor van de thesis én docent over dit onderwerp. Dankzij zijn expertise over het relatieve leeftijdseffect en zijn vakkundige kennisbijdrage heeft hij geholpen om deze thesis tot een mooi eindresultaat te brengen. We zijn hem ook dankbaar voor de opportuniteit om onze thesis als artikel te kunnen publiceren in het gewaardeerde 'Frontiers in Psychology' tijdschrift.

Hierbij willen we absoluut ook Dra. Sofie Bolckmans en prof. dr. em. Janet Starkes bedanken voor het reviewen van deze thesis en hun nuttige opmerkingen.
Evenzeer bedanken we het comité van het toelatingsexamen voor arts en tandarts die ons toegang gaven tot de data.

Ten slotte willen we ook onze respectieve familieleden bedanken voor hun morele ondersteuning in het volbrengen van deze thesis. Aurelia wenst haar partner zeer graag te bedanken voor zijn steun en geduld. Dina wenst specifiek haar twee zoontjes te bedanken voor hun liefdevolle geduld en emotioneel aanwezig zijn.

Mol, 26/03/2023 A. S.
Beverlo, 26/03/2023 D. D.

Situering

Deze masterproef werd volbracht aan de Faculteit Bewegings- en Revalidatiewetenschappen van de KU Leuven. Aan de Faculteit wordt reeds jarenlang onderzoek gedaan naar het relatief leeftijdseffect (RLE) onder leiding van prof. Helsen (5, 55)
Deze masterproef bestudeert het RLE in het toelatingsexamen voor geneeskunde, maar dan net specifiek voor wat betreft het nieuwe ingangsexamen. Vanaf 2018 bestaat de eindbeoordeling namelijk niet louter uit geslaagd of niet geslaagd, maar wordt men in één van de vijf categorieën gezet waarbij het cijfer van de categorie indicatief is voor het wel of niet succesvol geslaagd zijn voor het toelatingsexamen. Dit maakt de analyse een stuk complexer en extra interessant om (opnieuw) het RLE te onderzoeken bij het toelatingsexamen voor geneeskunde.

Het RLE is, behalve in de sport, immers ook aanwezig in het onderwijs. Vaak is er concurrentie, waarbij relatief oudere kinderen een voordeel hebben wat schoolprestatie betreft. Verschillen in de geboortemaand hebben ook een belangrijk effect op schoolprestaties. Dit is vooral het geval in de eerste jaren van de schoolloopbaan. Deze verschillen zijn een sleutelfactor bij mogelijke beslissingen over zowel de toelating tot het schoolsysteem als de herhaling van een leerjaar. Het RLE heeft een grotere impact naarmate de student jonger is. Tevens wordt vastgesteld dat deze (negatieve) effecten terugkeren op de leeftijd dat jongeren de overstap maken naar het hoger onderwijs omdat het RLE uitdooft tegen de leeftijd van 15-17 jaar en het is net vanaf deze 17-jarige leeftijd dat studenten deelnemen aan het toelatingsexamen voor arts/tandarts (23).

In onze studie wordt verondersteld dat het RLE ook gevolgen heeft voor studenten die deelnemen aan het Vlaamse toelatingsexamen geneeskunde en tandheelkunde. De gegevens werden geanalyseerd zoals verzameld en aangeleverd door het organisatiecomité van het Vlaams toelatingsexamen voor artsen en tandartsen. Gegevens van 22.990 studenten over vier jaar (2018-2021) werden geanalyseerd.

Uit de resultaten bleek dat: (1) er een duidelijk RLE is met oververtegenwoordiging van het aantal deelnemers geboren in kwartaal 1 en 2 en een duidelijke afname van deelnemers vanaf kwartaal 3; (2) er de voorbije vier academiejaren meer dan twee keer zoveel vrouwen zijn die deelnamen dan mannen; (3) er een omgekeerd RLE is wat betreft de slaagpercentages voor het toelatingsexamen met een duidelijke toename van het relatieve aantal toegelaten studenten geboren in het laatste kwartaal (4) mannelijke deelnemers een hoger slagingspercentage vertonen dan vrouwelijke deelnemers, vermoedelijk omdat meerkeuze-examenvragen met giscorrectie meer mannelijk risicogedrag bevorderen en resulteren in een voorzichtige antwoordstrategie bij vrouwelijke studenten. (56, 57, 61, 75) Het medische toelatingsexamen met meerkeuzevragen en giscorrectie is bovendien wellicht niet het meest geschikte hulpmiddel om het kennisniveau van studenten te meten, onder meer omdat het de capaciteiten van vrouwen onderschat, zoals ook in de literatuur reeds werd aangetoond. Het zou nuttig zijn voor onderzoekers, docenten en onderwijs beleidsmakers om zich meer bewust te zijn van het RLE. Het doel is om te evolueren naar toelatingsexamens voor arts/tandarts met eerlijkere slaagkansen voor die studenten die het gevoeligst zijn voor het RLE (merendeel vrouwelijk en de jongste onder leeftijdsgenoten). Daarom moeten interventies en mogelijke oplossingen worden overwogen en geëvalueerd.

Referenties

5. Helsen WF, van Winckel J, Williams AM. The relative age effect in youth soccer across Europe. J Sports Sci. 2005 Jun;23(6):629–636.

23. Smith AH. Born too late? How relative age affects college enrolment patterns. Research in Post-Compulsory Education. 2022 Oct 2;27(4):525–548.

55. Grondin S. Influence du trimestre de naissance sur l'acces à la participation au hockey et au volley-ball [microforme]. Thèse (M.Sc.)–Université de Sherbrooke; 1982.

56. Ben-Shakhar G, Sinai Y. Gender Differences in Multiple-Choice Tests: The Role of Differential Guessing Tendencies. J Educational Measurement. 1991 Mar;28(1):23–35.

57. Gender Gap and Multiple Choice Exams in Public Selection Processes [Internet]. [cited 2023 Mar 3]. Available from: https://ideas.repec.org/a/hpe/journl/y2020v235i4p11-28.html

61. Kelly S, Dennick R. Evidence of gender bias in True-False-Abstain medical examinations. BMC Med Educ. 2009 Jun 7;9:32.

75. Gender Bias in College Admissions Tests - Fairtest [Internet]. [cited 2023 Mar 3]. Available from: https://fairtest.org/gender-bias-college-admissions-tests/

Impact of the relative age effect and gender on the new entrance exam for medicine and dentistry education

Abstract

The aim of this retrospective study was to determine if: (1) there is an overrepresentation of students born in Q1 versus Q4 partaking in the entrance exam for medicine and dentistry, (2) there is an inverse relationship between RAE and birth quarter, (3) RAEs are still present when considering the success rates of the entrance exam, (4) whether there are gender differences in the success rate. The data from 22,990 students sitting for the entrance exam for medicine/dentistry (2018-2021) were analyzed. The results showed a clear RAE ($p<0.01$) (Q1 26.6%, Q4 22.7%). There was a reversal in RAE when considering the success rate, (Q1 24.4%, Q4 26.9%). Almost twice as many females than males participated, although they had a success rate of only 23.0% compared to 30.5% for males.

Key-words: Individual differences, birthdate, admission exam, university, gender inequality

Introduction

"Since one cannot choose one's parents, one can neither choose their month of birth". (sic)

A person's birth month may have a lifelong impact as it has been found to be closely related to success in physical and cognitive performance throughout the lifespan. Especially if we consider the Matthew effect of accumulated (dis)advantage. This effect refers to a pattern in which those who begin with an advantage accumulate more advantages over time and those who begin with a disadvantage become more disadvantaged over time (1,2).

Over the past decades, numerous studies have identified the impact of the relative age effect (RAE) on one's success in sports (3–5), chess (6,7) and academic achievement (8–10). Relative age is defined as the difference in age between two or more subjects within a date range of a year (11). Thus, the potential age difference between two classmates can be up to twelve months. The consequences are that the youngest can be expected to perform worse at the moment of entry, because they are in a later developmental stage, both physically and cognitively. These consequences are known as the relative age effect (12). Studies have shown that RAEs are both systematic and persistent in the school system (4,11). This effect is more important the younger the student is, but it reverses when reaching higher education.

In the Belgian school system, as in many other countries, all children of a specific age are admitted to school at a standardized time in the year. As a result, a group of children enter school at the same moment, regardless of their month of birth. Thus, the youngest children are almost one year younger than the oldest ones. In our study, we considered January 1st as the start of the selection year.

The earliest evidence of RAE was in the field of education. Studies revealed that RAEs are both systematic and pervasive in the school system, and that it also persists throughout the course of education (4,11,13–18). This is a phenomenon originally published by Pintner and Forlano in 1934, and by Huntington in 1938 (19,20). Likewise, for sports, it appears that being older confers advantages in readiness for intellectual development (21). For primary school children, the month of birth has a significant impact on all cognitive, motor and neuropsychological functions, such as attention, perception, or memory, as well as on those linked to process control and cognitive self-regulation. This is especially true during the first years of school (22). These differences appear to be a key factor in relation to possible decisions concerning both entry into the school system and grade repetition. The impact of age differences is notably greatest in the early years of school (23). Specifically, there is consistent evidence that in countries where school officially starts on September 1st (as in Belgium and other European and North American countries), children born in summer, perform lower in primary school than children born in the fall (24). Studies have shown that differences in month of birth have a significant effect on academic achievement. Younger students consistently score lower on tests of academic ability than their older peers throughout compulsory education (25,26). These differences are variously referred to in the literature as "birthday effect" (27), "birth month effect," "school starting age effect," (28,29), "birth season effect," "age effect," and/or "month of birth effect" (30–33). Relatively older children tend to score higher marks across subject areas, are more often enrolled in gifted and talented programs and are more likely to represent their school in various sports. Bjerke's findings in the numeracy field were consistent with Aune's findings from the Norwegian numeracy test. (14) One possible repercussion of RAE is that teachers expect less of their younger students due to their inherent poor academic and social performance which in turn is due to their later development. In this self-fulfilling prophecy, the lower expectations from teachers towards children who initially show a lesser degree of maturity, reinforces lower performance of these relatively younger children.

On the other hand, there is significant evidence from many countries around the globe (fig.7, PISA study 2020) suggesting that students who are relatively younger at school, tend to: (i) perform worse on achievement tests (34), (ii) are expected to struggle more or be held back a grade more often throughout elementary school (as high as 10% then declining in the latter years of education to 3-5%) (4,15). As a result, these students are less likely to attend college (35). In the three main domains assessed by PISA (Program for International Student Assessment) and on the student's progress through education, the youngest in their grade cohort at school entry were more likely to have repeated a grade in primary school (36). It was found that month of birth influenced success amongst 7-year-old children, and it even made a significant difference in success of children aged 11–14 (37). Mühlenweg and Puhani covered seventeen countries in their study and reached the conclusion that those who are younger when starting primary school experienced victimization in terms of psychological and personal development as well as academic characteristics (38). Likewise, in Givord (36) and Smith (23), the RAE is the most reliable predictor of differences in academic performance up to the age of 16 years. Finally, a British study found that fewer students than predicted, born between September and December, enrolled at college, while there were more students than expected, born between May and August (23).

However, it is in sports where the majority of research has been devoted to RAE since Grondin's historical work around 1984. This was the first time RAE was described in hockey and later in volleyball (55). The process that determined RAE in sports is equivalent to the one in education. RAE in educational formats seems to be enhanced by the fact that an individual will increase performance when expectations are greater, which suggests that children who perceive themselves capable of performing at a high level, and who think that they are talented, are more likely to continue perfecting their abilities and talents and invest more time and effort into school and other areas in life, with predictable results.

Rosenthal and Jacobsen (1968), among others, showed that teacher expectations impact on student performance (39). Positive expectations influence performance positively, and negative expectations influence performance negatively (39). As explained earlier, the Mathew effect plays a role as well. This effect refers to a pattern in which those who begin with an advantage accumulate more advantages over time and those who begin with a disadvantage become more disadvantaged over time (1,2). The result is ever widening of the differences between the advantaged and disadvantaged. There is no denying that the Matthew Effect is true for students. Students who are good readers experience more success, and they are encouraged by that success to read more (Pygmalion effect). As they read more, they become even more successful at reading. Their vocabulary and comprehension grow. Readers who struggle at decoding are less likely to want to pick up a book. They get much less practice and fall behind their peers. They fall behind in language arts classes and in content areas such as history and science. This effect has also been witnessed in soccer amongst the youngest players. Those who had the chance to practice more hours progressed significantly compared to the youngest of their peers. Later in life they still benefited from this. These effects are known as the Pygmalion effect and the Galatea effect, respectively. Put succinctly, the power of expectations cannot be overestimated.

The psychological consequences of this advantage may also create greater confidence and self-esteem derived from a comparison of ability to younger, cognitively less mature members of an age group (40). However, the reverse is alarmingly also true, relatively younger students are overrepresented in referral for psychiatric support, and generally display greater health problems (41,42). Relatively younger children achieve lower scores and more often have special needs, including special supportive education (15,43). They are also more likely to be diagnosed with a learning disability (35) and with attention-deficit/hyperactivity disorder (44).

Perhaps the most disturbing consequence of the RAE was described in 1999 when researchers discovered higher incidences of suicide in those born later in the year compared with their earlier born peers within school entry cohorts (45). An alarming study in 2015 examined how relative age in a grade effects suicide mortality rates of adolescents and young adults between 15 and 25 years of age (27). They associated suicide with the mental health of adolescents and young adults, who may be affected by their experiences at school and academic achievement (27). Apparently, students who struggle academically are more likely to suffer from depression and hopelessness, one of the major factors associated with suicide, than peers who succeed in academics (46–48). Other studies have documented a direct link between low academic performance and suicidal behaviors because of lower confidence and self-esteem (49–52). Neutralizing the negative effects of relative age should have important personal and social consequences on relatively younger children.

Fortunately, the size of RAEs in school tests has proven to be inversely correlated with age, as the relative age difference between children diminishes over time, as shown for cognitive abilities and performance (31,53,54). Current research suggests that the relative age effect dwindles by age 16 (37), however no research relating to the relative age effect is available for students aged 16–18 in higher education.

Because our study involves admission exam performance, it is also important to determine whether there is a relationship between gender and how people perform on tests that employ multiple choice vs open-ended questions. Studies of guessing tendencies and gender revealed that females tend to omit more items than males. This difference could not be attributed just to the better performance of males on most subtests because females omitted more items even on subtests which showed no significant differences in performance between the genders (56). Males answered questions faster and skipped more of them. Previous research supports the idea that girls can be less engaged with multiple-choice questions. Females tend to prefer questions that require more analysis and varied solutions, while males are more likely to just state their answers and show a lack of effort when there are more open-ended questions (57). The question remains whether this would strengthen or lessen the RAE and systematic gender bias, favoring males depending on the format of the entrance exam. The entrance exam for medicine and dentistry education is a multiple-choice exam which by its very nature tends to favor males.

Therefore, we hypothesized that RAE impacts students sitting the Flemish admission exam of medicine and dentistry. As far as we know, this is the very first study to examine this topic. Our aim was to determine if: (1) there is an overrepresentation of students born in the first quartile and an underrepresentation of students born in the last quartile between the students participating in the entrance exam of medicine and dentistry, (2) there is an inverse relationship between RAE and age category. Furthermore, we also wanted to determine if: (3) RAEs are still present when considering the success rates of the entrance exam, (4) students who were born in the first and second quartile have a significantly higher success rate, (5) whether there are gender differences in success rate.

Materials and methods

Participants

Many academic institutions are regulated by admission restrictions and conduct entrance exams to select students for a particular education. Passing the entrance exam is mandatory to be granted admission to medical education at one of the five Flemish and Brussels universities. Most students sit the exam at age 18, although a minority write it at a younger or older age for various reasons.

Data were collected and provided by the organizational committee of the Flemish entrance exam for medical doctors and dentists. Overall, 22,900 students (30.8% males and 69.2% females) participated in the entrance exam over the past four years (academic year 2017-2021). Data were analyzed per academic year. In case of similar results, the years were combined and regarded as one larger pool of data. We recorded the number of participating and successful students born in each month along with the student's gender. We examined birthdate through the pseudonymised national register number of the participants based on: (i) whether one already participated in a previous exam (in the same year, or years before), (ii) whether one meets the conditions to be ranked, (iii) and the success rate of each individual student. Using the birthdate of each individual student made it possible to control the results for just the 18-year-olds. There are two different entrance examination dates in the month of July, one for medicine and another for dentistry.

Entrance exam

As described above, previous literature clearly illustrates that RAE is present in education, team and individual sports. In these areas, there is often a tendency for strong competition, in which the relatively older children will have a relative advantage. The academic medicine and dentistry programs are very competitive as well. To become a candidate for medical specialization at the end of the 6-year program in Belgium, you need to be top of the class to escape the competitive screening and ranking system. Students are assessed on multiple factors such as their knowledge, practical internships, scientific work, etc., and ultimately only the best students are allowed to enter their training of choice. However, to start one's medical education it is necessary to write the competitive medical admissions test. Only a limited number of students per year will be admitted by the government to start medical schooling. We were interested in whether there might be a significant impact of RAE on the selection process.

Early 2018, a change in the exam requirements was implemented. Not all students who passed the entrance exam were permitted to start their medical studies as was the case before 2018. To follow medical or dentistry academic training at a Dutch-speaking university in Belgium, one must sit for a digital medical entrance exam, pass, and finally be ranked favorably. There are no specific conditions to participate in the exams. No restrictions exist in age nor the number of times you can participate. To pass the entrance exams three conditions must be met: (i) obtain at least a 50% score on the knowledge in sciences (KIS) component that mainly examines the competences in the basic sciences, with a focus on the subjects of biology, physics, chemistry and mathematics, (ii) obtain at least a 50% score on the general competences (GC) component that assesses general competences that are important, mainly the communicative competences (e.g. conflict management, listening skills, empathy, attention, reflection and respect) and a skill test which considers analyzing and reasoning skills. The third criterion (iii) is having graduated high school and obtained a diploma at the latest on September 30th of the same calendar year in which ranking for the entrance examination for medicine and dentistry medicine happen.

Rating and ranking

Since 2018, the total score plays a crucial role in one's ranking and therefore admission to the program. All questions are multiple choice with only one correct answer. The examination board applies a guess correction (three points per correct answer and minus one point per wrong answer). If one does not meet the three requirements, the candidate will not pass.

The final ranking is based on the results that are achieved. For the entrance examination for medicine, the first X participants (defined by numerus clausus) are ranked favorably (unless they have indicated at registration not to be ranked). We use X since this is a variable that varies per year, set by the committee who guard the numerus clausus. For the dental entrance exam, the first Y participants (defined by numerus clausus) are ranked favorably (unless they have indicated at registration not to be ranked). Non-ranking is for those that didn't comply with the third condition (having a diploma by the start of taking the admission exam). This means that they participate with the sole purpose of "trying out" the exams as a preparation for the consecutive year.

Also, since 2020 it has become impossible to be ranked favorably on both exams (medicine and dentistry). Candidates must choose the degree program when they register: medicine or dentistry. In that case, one cannot start with the other program even if they met the three conditions but fail to be ranked favorably.

Design and analysis

The birth months of the candidates were first divided into quartiles. The first quartile (Q1) starts from the beginning of January until the end of March (January–March) while the other quartiles are Q2 (April–June), Q3 (July-September), and Q4 (October–December) (Table 1). In fact, since the start of the selection year in football was shifted in 1997 from August 1st to January 1st, a school year and a football season both start in January in most of the European countries. In the schools in the UK, however, September 1st is still used as the start of a selection year, as they do in youth football, too.

To make an easy distinction, we also had the numbered ranking provided by the organization committee of the entrance exam for medicine and dentistry: 1: favorably ranked; 2: passed, but not ranked favorably due to result below cut-off point; 3: doesn't want to be ranked (regardless of result); 4: failed; 5: passed, but not ranked favorably because of favorably ranked for preferred education (result above cut-off mark).

A Spearman correlation was calculated between the month of birth and the participants per month of birth, between the month of birth and the percentage of participants that were admitted or failed. A chi square test was calculated as well. Statistical significance was set at $p < 0.01$.

Results

Participants per birth date quartile

The study population consisted of 22,900 participants, 15,905 females and 7,085 males, who participated in the entrance exam over the past four years. Results of the comparison between the participant's per birth month and birth quartile showed that there were notable differences in participation (Table 1). There is a clear decrease in participants from birth quartile 3 onwards. As can be seen from Table 1, birth quartiles 1 and 2 are overrepresented in the number of participants. More than twice as many females as males participated over the past four academic years (Table 1).

A negative correlation of -0.83 was found between the month of birth and the participants per month of birth, with a p-value of $p < 0.001$. The focus on birth quartiles (between the quartile of birth and the participants per quartile of birth) generates a negative correlation of -1.00 with a p-value of $p < 0.001$.

Table 1: Participants per birth quarter and per month of birth.

Birth quarter	**Participants per birth quarter**	**Percentage per birth quarter (%)**	**Month of birth**	**Participants per month of birth**	**Male**	**Female**
Q1	6106	26.6	January	2164	699	1465
			February	1945	619	1326
			March	1997	580	1417
Q2	6008	26.1	April	2001	614	1387
			May	2080	667	1413
			June	1927	550	1376
Q3	5649	24.6	July	2046	662	1384
			August	1805	551	1254
			September	1798	562	1236
Q4	5228	22.7	October	1804	522	1282
			November	1720	540	1180
			December	1704	519	1185
Total	**22990**	**100**		**22990**	**7085**	**15905**

Admitted to the program or not

Just a small portion of the participants were admitted to the program. Table 2 shows for each birth quartile the percentages of students who were admitted and those who failed.

Table 2: Percentage of participants per birth quarter and month that were admitted to the program and did not succeed

Birth quarter	Birth month	Percentage of participants that were admitted to the program (%)		Percentage of participants that did not succeed (%)	
	Jan		23.6		56.8
Q1	Feb	24.4	25.9	56.7	55.3
	Mar		23.7		58.1
	Apr		24.3		56.5
Q2	May	24.3	24.3	56.1	55.1
	Jun		24.1		56.8
	Jul		25.5		54.7
Q3	Aug	26.0	26.8	54.7	54.5
	Sep		25.9		54.9
	Oct		27.0		54.2
Q4	Nov	26.9	26.5	54.9	55.1
	Dec		27.2		55.6
Average		**25.3**	**25.3**	**55.7**	**55.7**

There was a clear increase in the relative number of participants that were admitted to the program from the first to the last birth quartile (Table 2). A significant correlation of 0.80 ($p < 0.002$) was observed.

There was a clear decrease in the relative number of participants that failed from the first to the last birth quartile (Table 2). A negative trend was observed ($r=-0.80$; $p < 0.083$).

Gender distribution

Over the four years, a total of 7,085 males (30.8%) and 15,905 females (69.2%) participated. Approximately, in every birth quartile a 1:2 ratio for males versus females was observed (Figure 1).

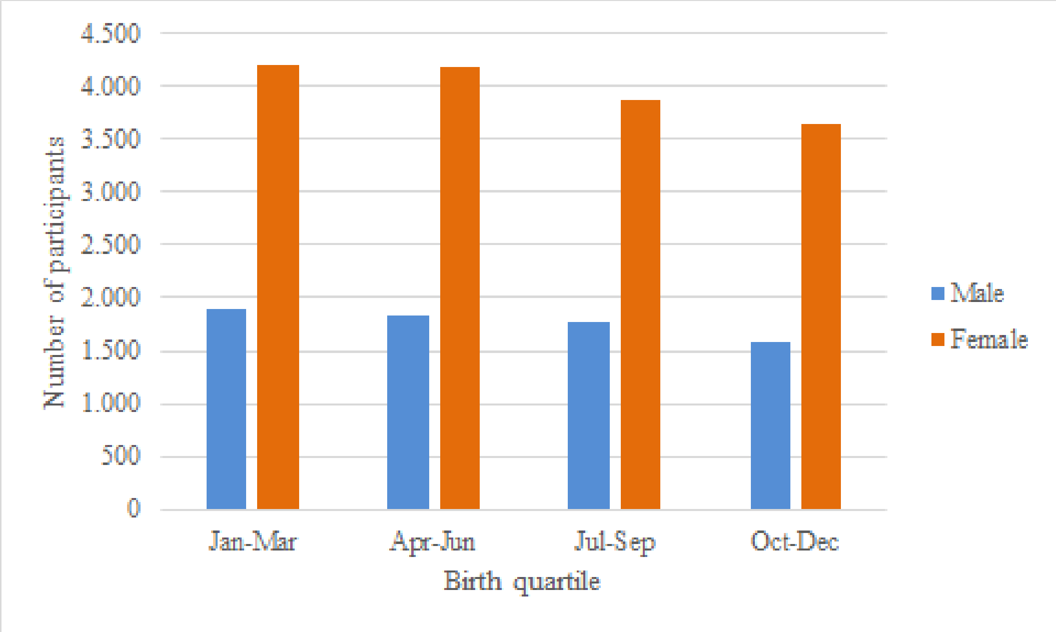

Figure 1: Gender distribution per birth quartile

Ranking

Participants were assigned a score from 1 to 5 depending on their ranking. 5,820 students (25.3%) received a 1, 1,181 students (5.1%) received a 2, 2,820 students (12.3%) a 3, 12,798 students (55.7%) a 4, and 372 students (1.6%) assigned a 5.

Therefore, the distribution of gender by category of decision was also examined (Table 3).

The results showed that 30.5% of the male participants were admitted to the program. 50.6% of them failed. For female participants, these percentages were 23.0% (who started their medical training) and 57.9% (who failed), respectively. This amounts to a total of 25.3% of the participating secondary school students who were eventually admitted to the medical training program and a total of 55.7% of all candidates who failed (Table 3).

Of all participants who were allowed to start their medical training, 37.2% were male and 62.8% were female. 28.0% of the students who failed the entrance exam were male and 72.0% were female. By comparison, 30.8% of the participants were male and 69.2% were female (Table 3).

Table 3: Gender and ranking.

Ranking	Absolute numbers			Percentage of ranking category per gender (%)			Percentage of gender in each ranking category (%)		
Gender	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**	**Male**	**Female**	**Total (%)**
1	2163	3657	5820	30.5	23	25.3	37.2	62.8	100
2	379	802	1181	5.4	5	5.1	32.1	67.9	100
3	847	1973	2820	12	12,4	12.3	30	70	100
4	3582	9215	12798	50.6	57.9	55.7	28	72	100
5	114	258	372	1.6	1.6	1.6	30.6	69.4	100
Total	**7085**	**15905**	**22990**	**100%**	**100%**	**100%**	**30.8**	**69.2**	**100**

Ranking: 1: favorably ranked; 2: passed, but not ranked favorably due to result below cut-off point; 3: doesn't want to be ranked (regardless of result); 4: failed; 5: passed, but not ranked favorably because of favorably ranked for preferred education (result above cut-off mark)

The focus of this study was whether there is a relative age effect present in the ranking categories of those who are admitted to the program and those who failed. Therefore, only these ranking categories (1 = admitted, 4 = failed) were examined in the tables (Table 4) and figures (Figure 2).

Table 4: Gender and ranking: percentage of gender in each ranking category.

Gender	Absolute			Percentage of gender in each ranking category (%) (not all categories are shown)	
Ranking	**Admitted**	**Failed**	**Total**	**Admitted**	**Failed**
Male	2163	3582	7085	37.2	28.0
Female	3657	9215	15905	62.8	72.0
Total	**5820**	**12798**	**22990**	**100**	**100**

37.2% of those who were admitted (group 1) were male, whereas 62.8% were female.

Only 28.0% of those who failed (group 4) were male, 72.0% were female (Table 4 and Figure 2). This is quite different from the aforementioned 1:2 ratio.

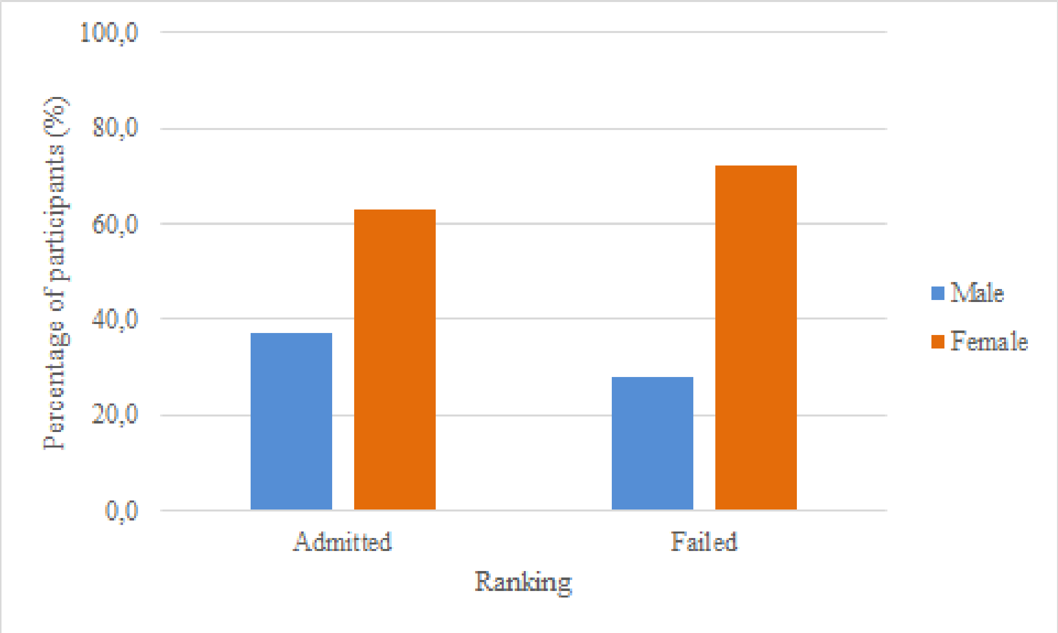

Figure 2: Percentage of participants per ranking and per gender.

The 1:2 ratio is visible in figure 2. For those who were admitted, this ratio seems to have remained the same. However, in those who failed we notice that the females were more than twice as likely to fail (Figure 2).

There is an overall success rate of 30.5% for male participants and 23.0% for female participants. We analyzed the distribution of this success rate per year as well. Male participants had a higher success rate than female participants. There was an increase in success rate from 2018 to 2021, for both male and female participants (Table 5).

Table 5: Success rate per year and per gender.

Success rate per year and gender	Male (%)	Female (%)
2018	27.7	19.7
2019	26.4	20.9
2020	34.6	27.0
2021	35.8	25.7

Discussion

Since 2018, the Flemish government adjusted the implementation of admission numbers for medicine and dentistry, with an innovative way of scoring and ranking the students. This is the very first study examining the impact of this new procedure.

Specifically, we analyzed if any RAE effect was present both in participating as well as passing rates for the medicine/dentistry entrance exam. We expected an overrepresentation of students born in the first quartile, compared to the last quartile. As well we hypothesized an inequality in students regarding gender, favoring males.

RAE and the entrance exam for the faculty of medicine

The RAE clearly impacts on the entrance exam for the faculty of medicine. The number of participants systematically decreased from the first to the last birth quartile.

RAE and pass rate

An observation of the number of participants (per birth quartile) who passed the entrance exam showed a reversed RAE effect. In a way, it is a positive phenomenon that the entrance exam has flattened out the RAE. If the RAE had remained, we might expect that relatively more students born in the first quartile would be admitted than those born in the last quartile. For the students who failed (ranking score 4), we might expect a greater percentage to be born in the last quartile. In both cases, we noticed the opposite, such that the RAE flattened out or was reversed. In fact, there was a positive correlation between month of birth and the number of students who passed ($r=0.80$; $p>0.002$), and a negative trend ($r=-0.80$; $p<0.083$) for those who failed. A potential explanation could be that Q4 students from the 4^{th} birth quartile happened to be better students who must show more effort and perseverance because they were exposed to a persistent relative disadvantage during childhood.

Gender distribution and pass rate

The past few years, there has been an obvious increase in female participants. The ratio is now around two to one (female:male). However, does this ratio remain if we look at who is admitted to the study of medicine/dentistry? And, equally important, are female participants still as representative when observing the pass rates? According to the last figure (figure 2), we observed that this ratio does stay the same. If we consider the absolute numbers (table 4), over the past four years, a total of 7,085 males (30.8%) participated and 15,905 females (69.2%). However, of the students who were admitted to the program, 37.2% were male and 62.8% were female. Of the students who didn't pass the entrance exam, 28% were male and 72% were female. Therefore, we can conclude that relatively more males then females passed the exam. A potential explanation might be the advantage of males in taking risks when it comes to multiple choice questions with guess correction (57).

Pass rate over the four years

The overall success rate generally increased from 2018 to 2021. This is probably due to a change in ranking categories over these years. The numbered ranking provided by the organization committee of the entrance exam for medicine and dentistry implemented: 1: favorably ranked, 2: passed, but not ranked favorably due to result below cut-off point, 3: doesn't want to be ranked (regardless of result), 4: failed, 5: passed, but not ranked favorably because of favorably ranked for preferred education (result above cut-off mark). In 2018 only ranking 1, 3 and 4 existed. In 2019, there was one more ranking (1, 2, 3 and 4), in 2020 only ranking 1, 2, 4 and 5, and in 2021 all five ranking categories were used. Therefore, the success rates may be slightly different over the four years.

Practical implications

Sweeney (2022) revealed how being an athlete born at the end of the year could be an advantage for long-term sport development due to overcoming adversities and demands derived from the RAE (e.g., they excel in technical qualities) to the most elite level of hockey play (58). This overrepresentation of athletes who are relatively younger or born in the last months of the year, is called the reversed RAE (15). In other words, young athletes keep fighting against their odds of expected RAE. This likewise supports the 'underdog' hypothesis, where relatively younger players are thought to benefit by more competitive play with their older counterparts. Also, Sweeney saw that the average career duration may be longer for players born later in the year (58). Similarly, the clear pattern of RAE as overrepresentation is not always reflected when one examines eventual earnings, skill, and performance.

When analyzing student admittance based on gender, we observed a greater proportion of females among the university students (69.2%). This could be partially explained by the existence of studies demonstrating that girls consistently outperform boys in scholastic attainment in primary and secondary studies. This results in a disproportionately greater number of females admitted to university. In addition, a study of gender, birthdates, and cognitive abilities of children with special educational needs showed that boys were significantly overrepresented compared to girls (59).

In previous population studies (60), differences in mean scores in numeracy and languages have been observed between boys and girls. In other studies, the examination of the relationship between guessing tendencies and gender revealed that females tend to omit more items than males, as the latter answered questions faster and guessed more of them (56). Across eight years of data for the first two years of an undergraduate medical curriculum, multiple-choice questions were found to contain the highest disparity between genders. Males were found to be 16.7 times more likely to outscore females in tests containing this item type. The proposed explanation for this large discrepancy is the higher propensity among females to abstain, with a corresponding result that males outscore females by three per cent. This trend extended to true/false questions in anatomy and physiology. Female advantage was recorded in one year in which short-response items were included. The common argument is that multiple-choice items favor greater male risk-taking behavior and more cautious response-solving strategies by female test takers (61). Hirschfeld, Moore and Brown (1995) cite ‘willingness to guess’ as a source of gender bias in multiple-choice item types (62). The authors argued that tests which do not penalize test takers for incorrect responses show less gender bias than those which do. The authors concluded that this was because males were more confident than females in high stakes testing environments. This general finding was replicated by Baldiga (2014) who found that in situations in which test takers are penalized for incorrect responses, females respond to significantly fewer questions than males (63). However, Baldiga found no difference between males and females in terms of confidence (a self-reported measure for each item) or in participants’ knowledge of the materials presented to them (United States and world history SAT). Given a male and a female with similar self-reported probabilities of getting a question correct, the female is one-third more likely to omit the question than the male if incorrect responses are penalized. This difference is explained via different risk preferences associated with the high-pressure environment, although Baldiga concedes that this accounts for only 40 per cent of the observed gender gap.

Previous research supports the idea that female students can be less engaged with multiple-choice questions. Girls tend to prefer questions that require more analysis and varied solutions while boys are more likely to just state their answers. Boys were more likely to show a lack of effort considering multiple choice questions than when there were more open-ended questions (57).

Overcoming the disadvantage of RAE in education – Possible Solutions

Interestingly, more than three decades ago, Bell and Daniels (1990) concluded: “Although it is difficult to see how schools could be organized so that there are no age differences in teaching groups, there should be research into mitigating this birthdate effect if the educational system is to be fair to all children” (24). Remarkably, in the last twenty years there hasn’t been much research regarding RAE in the educational system (69).

Classification according to month of birth is a much more effective practice than remedial school programs. In remedial school programs students from different class groups put together might still not have the same mindset, motivation, or skills. This can cause problems, especially for younger age groupings. To prevent these problems, Givord (2020) proposed that school systems should be shaped according to the month in which students were born (36). As a result, the disadvantages and deficiencies may be eliminated. Calsamiglia and Loviglio (2020) concluded that the school starting month is one of the most important criteria that determines success in education (70). They also discovered that younger students performed less well than older students and had higher rates of grade retention or drop-out. As well, they emphasized that this situation causes bigger problems in the following years. So, they advocated the view that some applications should be made to postpone the school starting age or reduce the birth month effect to prevent this negativity. However, delayed school entrance may cause different problems. The disadvantage may be eliminated for children whose starting school is delayed through postponement, but the problem persists for students who are not held back..

Jimerson (2001) also emphasized that the month of birth has effects not only on the academic success of children but also on their psycho-social status (71). As a result, as Campbell mentioned, the month of birth is closely related to success with cognitive and non-cognitive skills throughout childhood (22). The effect of the month of birth in children, who are disadvantaged due to developmental disabilities when starting school, deepens with the passing days, creating a kind of Matthew Effect and leading to failure and negativity in the following processes. In other words, the effect of the month of birth not only affects the academic success of the individual with the combination of the Matthew effect with the advancing process but also cascades to effects on many variables such as career choice, financial gain, and socio-psychological structure. In this context, to prevent these problems fundamentally, the "classification model of students in primary schools according to their month of birth and developmental level", which has been proven successful by the results of the study, is recommended. In this way, the loss of individuals and their being overwhelmed and unqualified in society may be prevented. On the other hand, a more balanced and healthy societal structure will be advanced.

Although university admissions tests underestimate females' abilities, there is no definitive answer as to what causes this bias. It appears that several factors contribute to the gender gap that are associated with sociocultural factors related to family environment: mainly the educational level of the mother (72), and a high high-school national test result (73). The harsh reality of bias against females in Japan became headline news in 2020 after Tokyo Medical University was caught rigging entrance exam scores to limit the number of female students (74).

Multiple-choice format

The American Educational Testing Service and the College Board concluded by a joint study that the multiple-choice format itself is biased against females. A variety of question types on Advanced Placement tests (like the SAT (a standardized test used for college admissions in the United States)) was tested and they found that the gender gap narrowed or disappeared on all types of questions (e.g., short answer, essay, constructed response) except multiple choice. The researchers conclude, "The better relative performance of females on constructed-response tests has important implications for high-stakes standardized testing. If both types of tests measure important education outcomes, equity concerns would dictate a mix of the two types of assessment instruments." A fairer approach would be to change the multiple-choice questions to open-ended questioning. Just asking multiple choice questions is not a fair way of examination. We are aware that this would take more work concerning the correction and scoring of the exam. This however cannot be the reason to systematically favor, and not in the least for an exam that will determine the future of these young students and eventually our healthcare system. A transitional measure could be to introduce exams with a mixture of multiple-choice questions and open-ended questions. Also, multiple-choice exams may not be the most appropriate tools to measure students' levels of knowledge. Substantial evidence exists that females approach problem-solving differently than males; they are more likely to work a problem out completely, to consider more than one possible answer, and to check their answers. Ironically, while these are desirable traits in school and in life, they work against females on an exam that is supposed to predict their ability to do academic work.

The role of guessing in multiple choice exams

The medicine entrance exam is scored with a guessing penalty, which deducts one-quarter point for every incorrect answer. Questions left blank are simply scored as zero. The intent of this policy is to make random guessing inadvisable. Although, since one or two answer choices can usually be eliminated as obviously incorrect, it is often in the test-taker's best interest to make an "educated guess". Research indicated that males were more likely to take risks on the test and guess when they did not know the answer;whereas, females tend to answer the question only if they are sure they are correct. The unwillingness to make educated guesses on this exam has been shown to have a significant negative impact on scores. Multiple choice (MC) tests are easier to correct than essay questions, but they do not solicit real knowledge from students. They can also summarize facts in such a way that distorts some facts. MC tests are somewhat better than true or false, but both suffer from the same weakness: a student can make a good grade by simply guessing.

"Speededness"

Another factor that contributes to the gender gap is the fast-paced, or "speeded" nature of the test. On some sections of the exam, students must answer as many as 35 questions (some of them requiring lengthy passages to read) in 30 minutes – an average of only 51 seconds per question. Substantial evidence exists that females approach problem-solving differently than males; they are more likely to work a problem out completely, to consider more than one-possible answer, and to check their answers. Once again, while these are desirable traits in school and in life, they work against females on an exam that is supposed to predict their ability to do academic work. Numerous studies have found that when the time constraint is lifted from the test, females' scores improve markedly, while males' scores remain the same or increase slightly. Un-timed administrations of the test still show a small score difference between males and females, suggesting that "speededness" is only one of several factors that bias the exam against young females.

The Test-Makers' Excuse

Test company officials have suggested that the gender gap is caused by the fact that more females take the tests than males. They argued that the larger pool of females includes more low- scoring students, which in turn reduces the average score for females.

In fact, research showed that controlling for these variables did not explain the gap. Today, as in Sharp's study from 1989, no evidence has been found that females' lower scores on SAT exams could be attributed to the larger number of females taking the exam. The conclusion is that the causes of the gap lay elsewhere than in the demographic makeup of the male and female testing populations. If the scoring gap were caused solely by the larger pool of females taking the exam, females should still attain the same percentage of high scores as males. In fact, the opposite is true: the gender gap is largest in the highest score ranges (75).

Ranking score system

Another approach could be to change the ranking system. We have observed that although systematically more females participated in the exam, they had a lower pass rate. This could be catered for by not dividing the whole group of participants over the different categories, but to do this proportionally to the number of females and males. In this way, the pass rate would be equated. This would influence ranking category 1; the size of this category depended on the number of males and females who participated, but this would be different for males and females (cfr. females impaired by multiple choice questions). Just as the admission numbers vary yearly depending on how many students can be admitted to the study of medicine/dentistry, they should apply this to genders, X females and Y males would be able to start.

Rosenthal effect

Since the problem of RAE manifests itself from a young age, it is important to take this on as soon as possible. It has been shown that girls perform better at school in classes/groups with only girls, than when they are mixed (76). Consequently, it can be useful to split up the classes in primary/secondary education for a couple of hours between boys and girls, to stimulate the latter. This also means that teachers and school boards must be well informed about this issue to tackle these needs with the right knowledge and logistical support, especially in schools where socio-economic demography is challenged.

Allocation method

Differentiating according to the allocation date of students in a class might be also effective. This method would be an interesting derivative of the allocation method based on the midway point of chronological and developmental birthdates of youth soccer players (67).

This method has recently been chosen as a method to eliminate the RAE effect amongst young football players. If this could be applied to young students as a way to allocate them to classes, then hopefully this would attenuate the RAE effect also in the educational system. (75). However, they also mentioned the importance of mental and psychological considerations of this reallocation (e.g., not playing with their friends anymore or youth players with a different cognitive maturation different from their physical development) (67).

Bolckmans (2022) examined whether this new allocation method eliminates the RAE. Arguably, participation in sport is not life-changing for most individuals; however, educational experiences and trajectories certainly are extremely significant. (68)

Conclusion

As children are separated into age groups based on chronological age, there are invariably cognitive, physical and maturational differences between the youngest and the oldest ones. Through high school, these effects remain present and decrease from late puberty age onwards (15-16 years). Despite this decrease, the effect remains until the end of academic education as we observed in our study. From our results, there is a clear RAE for the four years we considered. However, when we consider separately the years since 2018, we obtained different results.

Interestingly, with respect to gender, the number of male students stagnated while the number of female students showed a progressively increasing trend over the past decades. In 2017, there were twice as many female students as male students who sat the entrance exam and since 2018 even more females! Yet there is an overall success rate of 30.5% for male participants and 23.0% for female participants.

Beginning 2018, the entrance exam requirements for dentistry medicine changed. Not all students who passed the entrance exam were allowed to start medical education. Only the best 1,102 students were permitted to begin medical studies, because the number of participants is determined per year by the Flemish educational institution. Earlier results have shown a systemic 8% bias in favor of males, if only the best applicants are chosen this systemic bias favoring males is likely to increase in the new admission procedure (61).

It would be beneficial for researchers, teachers and education policymakers to be more aware of the RAEs, in order to find strategies to reduce attainment variations due to the relative age differences in any entrance exams. Therefore, interventions that seek to reduce RAEs and their consequences within schools and beyond should be considered and evaluated. Teachers should take into account birth-date when assessing the ability of pupils in their class, increasing the possibilities of people born in the last quartiles to enter the university (of medicine). In our opinion the entrance exam should be reevaluated, and several options have been discussed.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

In the context of our research, there was no contact with test subjects or human body material at any time. Our retrospective research study containing non-identifiable personal data was approved by an authorized medical ethics committee (Ethics committee of the Biomedical Sciences Group of KU Leuven, MP018905).

Author contributions

All authors were responsible for the conceptualization of the paper and the analysis and interpretation of the data. All authors contributed to manuscript revision and read and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

We thank the organization committee of the Flemish entrance exam for doctors and dentists for providing us the data.

References

1. O'Rand AM. The precious and the precocious: understanding cumulative disadvantage and cumulative advantage over the life course. Gerontologist. 1996 Apr;36(2):230–238.

2. Whitbourne SK, Dannefer WD. The "life drawing" as a measure of time perspective in adulthood. Int J Aging Hum Dev. 1985;22(2):147–155.

3. Delorme N, Boiché J, Raspaud M. The relative age effect in elite sport: the French case. Res Q Exerc Sport. 2009 Jun;80(2):336–344.

4. Cobley S, Baker J, Wattie N, McKenna J. Annual Age-Grouping and Athlete Development. Sports Med. 2009;39(3):235–256.

5. Helsen WF, van Winckel J, Williams AM. The relative age effect in youth soccer across Europe. J Sports Sci. 2005 Jun;23(6):629–636.

6. Breznik K, Law KMY. Relative Age Effect in Mind Games: The Evidence from Elite Chess. Percept Mot Skills. 2016 Apr;122(2):583–594.

7. Helsen WF, Baker J, Schorer J, Steingröver C, Wattie N, Starkes JL. Relative age effects in a cognitive task: A case study of youth chess. High Ability Studies. 2016 Jul 2;27(2):211–221.

8. Navarro J-J, García-Rubio J, Olivares PR. The relative age effect and its influence on academic performance. PLoS One. 2015 Oct 30;10(10):e0141895.

9. Yarim MA, Cemaloğlu N. The effect of date of birth on success: proposing a model for primary schools in the context of the Matthew effect. Reading & Writing Quarterly. 2022 Sep 9;1–22.

10. Sharp C. What's age got to do with it? A study of patterns of school entry and the impact of season of birth on school attainment. Educational Research. 1995 Dec;37(3):251–265.

11. Aune TK, Pedersen AV, Ingvaldsen RP, Dalen T. Relative age effect and gender differences in physical education attainment in norwegian schoolchildren. Scandinavian Journal of Educational Research. 2017 May 4;61(3):369–375.

12. Barnsley R, Thompson A, Barnsley PE. Hockey success and birthdate: The relative age effect. Canadian Association for Health, Physical Education, and Recreation. 1985 Jan 1;51:23–28.

13. Dalen T, Ingvaldsen RP, Roaas TV, Pedersen AV, Steen I, Aune TK. The impact of physical growth and relative age effect on assessment in physical education. Eur J Sport Sci. 2017 May;17(4):482–487.

14. Aune TK, Ingvaldsen RP, Vestheim OP, Bjerkeset O, Dalen T. Relative age effects and gender differences in the national test of numeracy: A population study of Norwegian children. Front Psychol. 2018 Jul 2(9):1091.

15. Cobley S, McKenna J, Baker J, Wattie N. How pervasive are relative age effects in secondary school education? J Educ Psychol. 2009 May;101(2):520–528.

16. Jinks PC. An investigation into the effect of date of birth on subsequent school performance. Educational Research. 1964 Jun;6(3):220–225.

17. Massey A, Elliott G, Ross E. Season of birth, sex and success in GCSE English, mathematics and science: some long-lasting effects from the early years? Research Papers in Education. 1996 Jun;11(2):129–150.

18. Sharp C, Hutchison D, Whetton C. How do season of birth and length of schooling affect children's attainment at key stage 1? Educational Research. 1994 Jun;36(2):107–121.

19. Pintner R, Forlano G. The birth month of eminent men. Journal of Applied Psychology. 1934 Apr;18(2):178–188.

20. Sandon F. Season of birth: its relation to human abilities. Eugen Rev. 1939 Jul;31(2):130–2.

21. Bell JF, Massey A, Dexter T. Birthdate and ratings of sporting achievement: analysis of physical education GCSE results. European Journal of Physical Education. 1997 Jan;2(2):160–166.

22. Campbell T. Relative age and the Early Years Foundation Stage Profile: How do birth month and peer group age composition determine attribution of a "Good Level of Development"—and what does this tell us about how "good" the Early Years Foundation Stage Profile is? Br Educ Res J. 2022 Apr;48(2):371–401.

23. Smith AH. Born too late? How relative age affects college enrolment patterns. Research in Post-Compulsory Education. 2022 Oct 2;27(4):525–548.

24. Bell JF, Daniels S. Are summer-born children disadvantaged? the birthdate effect in education. Oxford Review of Education. 1990 Jan;16(1):67–80.

25. McEwan PJ, Shapiro JS. The benefits of delayed primary school enrollment. J Human Resources. 2008;43(1):1–29.

26. Daniels K. Measures of five aspects of affective well-being at work. Human Relations. 2000 Feb 1;53(2):275–294.

27. Matsubayashi T, Ueda M. Relative Age in School and Suicide among Young Individuals in Japan: A Regression Discontinuity Approach. PLoS One. 2015 Aug 26;10(8):e0135349.

28. Black SE, Devereux PJ, Salvanes KG. Too young to leave the nest? The effects of school starting age. Review of Economics and Statistics. 2011 May;93(2):455–467.

29. Fredriksson P, Ockert B. Is early learning really more productive? the effect of school starting age on school and labor market performance. SSRN Journal. 2005;

30. Angrist JD, Keueger AB. Does compulsory school attendance affect schooling and earnings? Q J Econ. 1991 Nov 1;106(4):979–1014.

31. Musch J, Grondin S. Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport. Developmental Review. 2001 Jun;21(2):147–167.

32. Ponzo M, Scoppa V. The long-lasting effects of school entry age: Evidence from Italian students. Journal of Policy Modeling. 2014 May;36(3):578–599.

33. Robertson E. The effects of quarter of birth on academic outcomes at the elementary school level. Econ Educ Rev. 2011 Apr;30(2):300–311.

34. Kawaguchi D. Actual age at school entry, educational outcomes, and earnings. J Jpn Int Econ. 2011 Jun;25(2):64–80.

35. Dhuey E, Lipscomb S. What makes a leader? Relative age and high school leadership. Econ Educ Rev. 2008 Apr;27(2):173–183.

36. How student's month of birth is linked to performance at school: New evidence from PISA. 2020 May.

37. Crawford C, Dearden L, Greaves E. The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. J R Stat Soc Ser A, Stat Soc. 2014 Oct;177(4):829–860.

38. Mühlenweg AM, Puhani PA. The Evolution of the School-Entry Age Effect in a School Tracking System. J Human Resources. 2010;45(2):407–438.

39. Rosenthal R, Jacobson L. Pygmalion in the classroom. Urban Rev. 1968 Sep;3(1):16–20.

40. Cobley S, Abraham C, Baker J. Relative age effects on physical education attainment and school sport representation. Physical Education & Sport Pedagogy. 2008 Jul;13(3):267–276.

41. Goodman R, Gledhill J, Ford T. Child psychiatric disorder and relative age within school year: cross sectional survey of large population sample. BMJ. 2003 Aug 30;327(7413):472.

42. Elder TE, Lubotsky DH. Kindergarten Entrance Age and Children's Achievement. J Human Resources. 2009;44(3):641–683.

43. Wallingford EL, Prout HT. The relationship of season of birth and special education referral. Psychol Sch. 2000 Jul;37(4):379–387.

44. Zoëga H, Valdimarsdóttir UA, Hernández-Díaz S. Age, academic performance, and stimulant prescribing for ADHD: a nationwide cohort study. Pediatrics. 2012 Dec;130(6):1012–1018.

45. Thompson AH, Barnsley RH, Dyck RJ. A new factor in youth suicide: the relative age effect. Can J Psychiatry. 1999 Feb;44(1):82–85.

46. Bandura A, Pastorelli C, Barbaranelli C, Caprara GV. Self-efficacy pathways to childhood depression. J Pers Soc Psychol. 1999 Feb;76(2):258–269.

47. Quiroga CV, Janosz M, Bisset S, Morin AJS. Early adolescent depression symptoms and school dropout: Mediating processes involving self-reported academic competence and achievement. J Educ Psychol. 2013 May;105(2):552–560.

48. Fröjd SA, Nissinen ES, Pelkonen MUI, Marttunen MJ, Koivisto A-M, Kaltiala-Heino R. Depression and school performance in middle adolescent boys and girls. J Adolesc. 2008 Aug;31(4):485–498.

49. Buzzai C, Sorrenti L, Costa S, Toffle ME, Filippello P. The relationship between school-basic psychological need satisfaction and frustration, academic engagement and academic achievement. Sch Psychol Int. 2021 Oct;42(5):497–519.

50. Madge N, Hawton K, McMahon EM, Corcoran P, De Leo D, de Wilde EJ, et al. Psychological characteristics, stressful life events and deliberate self-harm: findings from the Child & Adolescent Self-harm in Europe (CASE) Study. Eur Child Adolesc Psychiatry. 2011 Oct;20(10):499–508.

51. Martin RP, Foels P, Clanton G, Moon K. Season of birth is related to child retention rates, achievement, and rate of diagnosis of specific LD. J Learn Disabil. 2004;37(4):307–317.

52. Taliaferro LA, Muehlenkamp JJ. Factors associated with current versus lifetime self-injury among high school and college students. Suicide Life Threat Behav. 2015 Feb;45(1):84–97.

53. Bisanz J, Morrison FJ, Dunn M. Effects of age and schooling on the acquisition of elementary quantitative skills. Dev Psychol. 1995 Mar;31(2):221–236.

54. Morrison FJ, Smith L, Dow-Ehrensberger M. Education and cognitive development: A natural experiment. Dev Psychol. 1995;31(5):789–799.

55. Grondin S. Influence du trimestre de naissance sur l'acces à la participation au hockey et au volley-ball [microforme]. Thèse (M.Sc.)–Université de Sherbrooke; 1982.

56. Ben-Shakhar G, Sinai Y. Gender Differences in Multiple-Choice Tests: The Role of Differential Guessing Tendencies. J Educational Measurement. 1991 Mar;28(1):23–35.

57. Gender Gap and Multiple Choice Exams in Public Selection Processes [Internet]. [cited 2023 Mar 3]. Available from: https://ideas.repec.org/a/hpe/journl/y2020v235i4p11-28.html

58. Sweeney L, Cumming SP, MacNamara Á, Horan D. A tale of two selection biases: The independent effects of relative age and biological maturity on player selection in the Football Association of Ireland's national talent pathway. Int J Sports Sci Coach. 2022 Sep 26;174795412211261.

59. Wilson G. The effects of season of birth, sex and cognitive abilities on the assessment of special educational needs. Educational Psychology. 2000 Jun;20(2):153–166.

60. Stoet G, Geary DC. Sex differences in mathematics and reading achievement are inversely related: within- and across-nation assessment of 10 years of PISA data. PLoS One. 2013 Mar 13;8(3):e57988.

61. Kelly S, Dennick R. Evidence of gender bias in True-False-Abstain medical examinations. BMC Med Educ. 2009 Jun 7;9:32.

62. Hirschfeld M, Moore RL, Brown E. Exploring the gender gap on the GRE subject test in economics. J Econ Educ. 1995 Jan;26(1):3–15.

63. Baldiga K. Gender differences in willingness to guess. Manage Sci. 2014 Feb;60(2):434–448.

64. Grossmann B, Lames M. Relative age effect (RAE) in football talents – the role of youth academies in transition to professional status in germany. International Journal of Performance Analysis in Sport. 2013 Apr;13(1):120–134.

65. Helsen WF, Baker J, Michiels S, Schorer J, Van Winckel J, Williams AM. The relative age effect in European professional soccer: did ten years of research make any difference? J Sports Sci. 2012 Sep 24;30(15):1665–1671.

66. Hill B, Sotiriadou P. Coach decision-making and the relative age effect on talent selection in football. European Sport Management Quarterly. 2016 May 26;16(3):292–315.

67. Helsen WF, Thomis M, Starkes JL, Vrijens S, Ooms G, MacMaster C, et al. Leveling the Playing Field: A New Proposed Method to Address Relative Age- and Maturity-Related Bias in Soccer. Front Sports Act Living. 2021 Mar 4;3:635379.

68. Bolckmans S, Starkes JL, Towlson C, Barnes C, Parkin G, Helsen WF. Leveling the Playing Field: A New Proposed Method to Address Relative Age- and Maturity-Related Bias in UK Male Academy Soccer Players. Front Sports Act Living. 2022 Mar 3;4:847438.

69. Urruticoechea A, Oliveri A, Vernazza E, Giménez-Dasí M, Martínez-Arias R, Martín-Babarro J. The relative age effects in educational development: A systematic review. Int J Environ Res Public Health. 2021 Aug 26;18(17).

70. Calsamiglia C, Loviglio A. Maturity and school outcomes in an inflexible system: evidence from Catalonia. SERIEs (Berl). 2020;11(1):1–49.

71. Jimerson SR. Meta-analysis of Grade Retention Research: Implications for Practice in the 21st Century. School Psych Rev. 2001 Sep 1;30(3):420–437.

72. Kong F, Chen Z, Xue S, Wang X, Liu J. Mother's but not father's education predicts general fluid intelligence in emerging adulthood: Behavioral and neuroanatomical evidence. Hum Brain Mapp. 2015 Nov;36(11):4582–4591.

73. Krings R, Huwendiek S, Walsh N, Stricker D, Berendonk C. Predictive power of high school educational attainment and the medical aptitude test for performance during the Bachelor program in human medicine at the University of Bern: a cohort study. Swiss Med Wkly. 2020 Dec 14;150:w20389.

74. Gender gap closes in medical school entrance exams - The Japan News [Internet]. [cited 2023 Mar 3]. Available from: https://japannews.yomiuri.co.jp/society/general-news/20220429-24031/

75. Gender Bias in College Admissions Tests - Fairtest [Internet]. [cited 2023 Mar 3]. Available from: https://fairtest.org/gender-bias-college-admissions-tests/

76. Pahlke E, Hyde JS, Allison CM. The effects of single-sex compared with coeducational schooling on students' performance and attitudes: a meta-analysis. Psychol Bull. 2014 Jul;140(4):1042–1072.