STUDENT ESSAY COMPETITION
WINNER, 2011

Relative age effects in education and sport: An argument for human, not statistical solutions

Nick Wattie

What a difference a date makes. Although seemingly harmless, a child’s date of birth may have a significant influence on their subsequent development. And no, this has nothing to do with celestial bodies and astrology, or even seasonal climactic factors such as temperature or sunlight exposure (see Wattie et al., 2008b). Rather, birth date can influence development because the social policies which govern primary and secondary education, and youth sport participation, use date of birth to determine eligibility for inclusion and to structure cohorts. For example, education and youth sport systems in England employ annual age grouping policies whereby children are grouped into cohorts using a selection date of September 1st. More specifically, to enter Year one of primary school a child must turn 6 years of age before September 1st of the current academic year. Therefore, a child born on August 31st would be placed in the same annual age group (aka cohort) as a child born on September 1st of the previous calendar year; making the child born at the end of August approximately 12 months younger that the child born at the beginning of September. Relative age is the term used to describe the aforementioned difference in age within an annual age group, and relative age effects (RAEs: Barmsley et al., 1985) is the term used to describe outcomes associated with relative age differences. Although the specific selection dates utilized may be different, e.g.
September 1\textsuperscript{st} vs. January 1\textsuperscript{st}, annual age grouping policies and the use of selection dates are internationally ubiquitous in both education (Bedard and Dhuey, 2006), and youth sport (Musch and Grondin, 2001). The notable consequence of relative age differences in education and sport is that relatively younger youth are more likely to be placed at a disadvantage compared to their relatively older peers.

The purpose of this essay is two-fold. First, a brief summary of RA\AE s in education and youth sport will be presented with the hope of spurring interest and attention to relative age phenomena among the Radical Statistics membership. Far from an exhaustive review, this section aims to provide a taste of the relative age literature. Second, a summary of the proposed solutions to RA\AE s will be presented, and an argument will be made that statistical issues exist in the literature that may influence the effectiveness of some solutions. Furthermore, this section argues that some proposed solutions to RA\AE s may, in fact, miss the mark.

In primary and secondary stages of education, relatively older pupils are more likely to attain higher grades than their relatively younger peers across a range of different subjects (Bell and Daniels, 1990, Massey et al., 1996, Sharp et al., 1994, Cobley et al., 2009b). The difference in attainment between relatively older and relatively younger pupils has been reported to be approximately 10\% at the beginning of secondary school, which generally decreases to about 3-5\% by the end of secondary school (Massey et al., 1996, Foxman et al., 1990). Standardized test scores from the 1995 and 1999 Trends in International Mathematics and Science Study of 19 OECD (www.oecd.org) countries revealed a 4-12 percentile disadvantage for the relatively youngest among 4\textsuperscript{th} grade children (9 years of age) and a 2-9 percentile difference for relatively younger 8\textsuperscript{th} grade children (13 years of age) (Bedard and Dhuey, 2006).

Where academic ability streaming exists - defined as the streaming, selection or organization of youth into different hierarchies of proficiency – relatively older pupils are more likely to be placed in top ability streams (Thompson, 1971) and to be identified as ‘gifted and talented’ (Cobley et al., 2009b). Perhaps not surprisingly, the other side of this coin is that relatively younger children may be more likely to be identified as having ‘special educational needs’ (Wilson, 2000, Bookbinder, 1967, Martin et al., 2004, Wallingford and Prout, 2000). Relatively younger pupils are also over-represented among those identified as learning disabled (Maddux, 1980) and those referred for psychological counselling due to academic and/or behavioural problems (Drabman et al., 1987, Tarnowski et al., 1990). Among
samples of pupils from England (Cobley et al., 2009b) and Wales (Carroll, 1992) relatively younger pupils had lower attendance rates. Indeed, research suggests that relatively younger pupils have lower levels of self-esteem (Thompson et al., 2004, Fenzel, 1992), although ambiguous results have also been reported for motivational and engagement outcomes (Martin, 2009).

Findings suggest that there might be some even more alarming outcomes associated with younger relative age. In a nationally representative sample of 5 to 15 year olds from the United Kingdom relatively younger pupils were found to be at greater risk of psychiatric disorders than relatively older pupils (Goodman et al., 2003) – with the authors suggesting that if the prevalence of psychiatric disorders were equal across different relative ages there could be approximately 60,000 fewer cases of child psychiatric disorder. Lastly, while examining deaths by suicide among individuals under the age of 20 within the province of Alberta, Canada, Thompson and colleagues (Thompson et al., 1999) observed an over-representation of relatively younger youths.

While the relative age literature has its origins in education-related research, sport-related relative age research has been rapidly expanding in recent decades (Wattie et al., 2008a). Like outcomes in education relatively older youth predominantly experience advantages over their younger peers. For example, research has shown that relatively older pupils receive higher grades in Physical Education (PE) classes (Cobley et al., 2008, Bell et al., 1997), and are more likely to be selected to school sports teams (Wilson, 1999, Cobley et al., 2008). However, the most notable RAEs emerge in competitive youth sport, particularly within ice hockey and soccer. Evidence in ice hockey and soccer suggests that there are 20% more youth than expected (compared to population statistics) among those born within the first three months of the selection year, i.e. the relatively oldest youth (Barnsley and Thompson, 1988, Barnsley et al., 1992, Grondin et al., 1984). Conversely, the same studies suggest that there are approximately 20% less youth than expected among those born the in the last three months of the selection, i.e., the relatively youngest. These RAEs emerge among youth as young as 8-9 years of age. Like education, ability streaming from a young age appears to be conducive to fostering RAEs in sport youth sport (Grondin et al., 1984, Musch and Grondin, 2001). What is more, RAEs persist into elite adult (professional) levels of play in ice hockey (Grondin and Trudeau, 1991, Montelpare et al., 1998, Wattie et al., 2007). Given the early emergence of RAEs in youth sport and its persistence into professional levels, Barnsley and Thompson (1988, p.175) quipped that,
“professional hockey players are really drafted when they are nine years old, at the time when they are selected for top tier leagues in their age group.”

The same trends that exist in ice hockey also exist in youth and elite level soccer. While most of the research on ice hockey has been with respect to Canadian youth and elite adult athletes, the evidence of RAEs in youth and elite level soccer has been truly international: RAEs have been identified among youth and elite soccer players in France, England, Germany, Sweden, Netherlands, Brazil, Japan, Australia, Spain, United States (Carling et al., 2009, Edgar and O’Donoghue, 2004, Dudink, 1994, Brewer et al., 1995, Musch, 2002, Simmons and Paull, 2001, Jimenez and Pain, 2008, Vincent and Glamser, 2006, Mujika et al., 2009, Helsen et al., 2005).

In addition to ice hockey and soccer, over-representations of relatively older players have also been documented for baseball in the United States (Thompson et al., 1991, Côté et al., 2006) and Japan (Grondin and Koren, 2000). Researchers have also found RAEs in rugby union in Australia (Abernethy and Farrow, 2005). More recently, one of the largest magnitude RAE was observed in competitive youth rugby league in England, with nearly 60% of all players being born in the first 3 months of the selection year (Till et al., 2010). Studies have even reported RAEs among professional race car (NASCAR) drivers (Abel and Kruger, 2007), and shooting sports (Delorme and Raspaud, 2009).

Allen and Barnsley eloquently and succinctly summarize the why of RAEs: “errors result from the difficulty, or impossibility, of observing ability independent of maturity in children.” (1993, p. 649). Simply put, the increased likelihood that relatively older youth are of advanced maturity, cognitive and/or physical, affords them a probabilistic (not deterministic) advantage over relatively younger youth. Other factors are important too. For example, in sport the cultural popularity of a sport and the degree of physicality inherent to a sport have been implicated as potential catalysts of RAEs (for a review see Musch and Grondin, 2001).

In addition to the relationship between ability and maturation, there are indications that a more insidious aetiology could be at work as well. For example, one interesting thing about the studies where relatively younger youth were more likely to be referred for counselling/behavioural problems, is that their scores on standardized tests of behavioural problems were equivalent to those of relatively older youth (Drabman et al., 1987). While this outcome might still be linked to behavioural maturation, it suggests that
inaccurate subjective perceptions by teachers, rather than objective differences in behaviour. There may be a need to discuss the possibility that some RAEs, such as lower attendance rates and referral for behavioural problems, are the pathologization of individuals’ normal reactions/behaviours within their developmental environment.

**Proposed Solutions for RAEs**

The suggested directions for eliminating RAEs can be collectively termed *technical solutions*. The technical solutions are comprised of i) selection criteria and grouping solutions and ii) handicap solutions (handicap solutions thus far being exclusive to education). The first of the *technical solutions* are selection criteria and grouping solutions. Selection criteria solutions primarily involve modification to annual age grouping procedures. Whereas most systems (sport and education) utilize a single selection date (e.g., September 1\textsuperscript{st}) to signify the beginning of the annual age group, these proposed solutions would involve rotating the selection to different points of the year. Some of these proposed solutions have been simple, only involving relatively straightforward changes to selection dates every year or bi-annually (e.g., the Novem system: Boucher and Halliwell, 1991), or the establishment of relative age quotas on sports teams (i.e., ensuring an equal distribution of youth in each relative age Quartile). Other selection criteria solutions, such as the Relative Age Fair (RAF) system (Hurley et al., 2001, Hurley, 2009), are far more computationally intense, although possibly more equitable for it. However, these different selection criteria solutions ultimately have the same goal: to ensure that no child is *always* advantaged or disadvantaged by relative age and they experience being different relative ages (sometimes older than their peers; sometimes younger), and/or to give equal treatment to youth of all relative age.

The second proposed *technical solution* for RAEs has been to implement what is essentially a handicap system, whereby the average difference in attainment associated with relative age would be adjusted so that no difference exists between relatively older and younger youth (e.g., age normalisation of test results: Crawford et al., 2007). Namely, the grades of relatively younger youth would be inflated so that any
probabilistic (dis)advantage conferred by relative age would be nullified.¹

While the proposed solutions to RAEs offer many potential avenues for intervention-based research, and may prove to be legitimate means to eliminate RAEs, some facets of the proposed solutions may be problematic. The commonality of the proposed technical solutions to RAEs has been to associate the irreducible technical origins of RAEs, i.e. the cut-off date of annual age grouping and/or relative age itself, with the conclusion that solutions must also be technical in nature. While the details and focus of each solution differ, their commonality rests in their technical focus. As such, there is the possibility that some RAEs may belong to what Hardin (1968, p.1243)² described as a “class of human problems which can be called no technical solution problems.” In Hardin’s seminal 1968 paper, he presented the thesis that for some problems (“no technical solution problems”) the solution will not be technical, it will be human. There is the possibility that some RAEs may be no technical solution problems.

The first issue concerns how relative age has been treated as a variable. While some points regarding the statistics of RAEs are either seldom discussed they are essential for an objective and accurate understanding of RAEs, and therefore essential to discussions of interventions and policy change. Rather than keeping relative age as a continuous variable, from 1 to 365 days, researchers have predominantly dummy coded relative age into a 4-level ordinal categorical variable, i.e. “Quartiles”. For example, Grondin et al.’s (1984) youth ice hockey populations were subject to an annual age grouping policy with a selection year beginning on January 1st, therefore all players with birthdates in January, February and March would have been coded as Quartile 1 (Q1; subsequently, April, May and June = Q2; July, August, September = Q3; October, November and December = Q4).³ Those born in Q1 are the relatively oldest, and those born Q4 the relatively youngest. Figure 1 displays fictitious, although representative depiction of how RAEs have been presented in the literature. While this method has advantages for dissemination –

¹ This practice is currently being used in selective non-fee paying grammar schools within the UK (see: www.nfer.ac.uk/nfer/research/assessment/eleven-plus/age-standardisation.cfm).

² Hardin was echoing and elaborating upon a thesis previously presented by Wiesner and York (1964; Scientific American, vol. 211).

³ Some exceptions exist of course, with relative age coded as a two-level (Edwards, 1995) or a three-level categorical variable (Simmons and Paull, 2001; Brewer et al., 1995).
it is visually friendly and easy to describe – it is not without limitations. Compared to keeping relative age as a continuous variable dummy coding it into a categorical variable is inherently information poor.\(^4\)

**Figure 1. An example of a typical RAE using the Quartile method.**

This method also assumes that being born on September 1\(^{st}\) is equivalent to being born on November 30\(^{th}\) (i.e., Quartile 1), but someone born on November 30\(^{th}\) (i.e., Quartile 1) is somewhat arbitrarily seen as being different than someone born on December 1\(^{st}\) (i.e., Quartile 2). Therefore, in some instances a day (or month) may mean nothing, but in others may be quantitatively different. In summary, the method of describing relative age in quartiles is somewhat arbitrary and dilutes the information inherent to relative age as a continuous parametric form of data.

In addition, some unexpected attainment trends have been reported when examining RAEs, suggesting that simple mean trends may obfuscate the complexity of relative age phenomena. For example, in an investigation of secondary school math exam results, Allen (2008) observed that relatively older pupils had higher average attainment in math. Nothing new there. However, Allen thought to look beyond the average, and found that relatively younger pupils were more likely (p < .05) than their relatively older counterparts to score above the 90\(^{th}\) percentile. On average relatively older pupils had higher grades, but

---

\(^4\) Whether or not this loss of information is consequential will, of course, depend on the empirical and explanatory context.
relatively younger pupils were more likely to have the highest grades. In essence, by solely describing RAEs using comparisons of aggregate statistics between Quartiles of birth, researchers risk committing an ecological fallacy because they are inferring that aggregate statistics apply to relationships at the individual level (Everitt, 2004). The results reported by Allen also stress the importance of remembering that advantages for relatively older youths are probabilistic and not deterministic. The effect sizes of RAEs are generally small to medium and as such there are a sizeable number of relatively younger children who do not succumb to the potential stricture of relative age.

The fact that the influence of relative age is probabilistic and not deterministic creates numerous potential problems for implementing relative age-related interventions and policy changes. For example, how does a technical relative age intervention, such as age normalisation of test results, influence relatively younger youth who are performing on-par with their relatively older counterparts? What if RAEs are evident among boys but not girls, as is often the case in sport (see Cobley et al., 2009a)? Would it be prudent to implement across all youth within a school or sport? The reality is that there may be a need for contextualized policy (Thrupp and Lupton, 2006): the choice and efficacy of interventions may be individual-specific and/or context specific.

There are also potential problems with technical solutions that are independent of issues with the use of Quartiles in research and solutions. While it is true that RAEs cannot exist without relative age per se, previous research findings suggest that modifications to grouping procedures or evaluation criteria (a handicap system) may be problematic for two reasons. First, RAEs may not be etiologically as simple as relative age differences or a single selection date, therefore none of these solutions might actually address many of the important factors that create and propagate RAEs. For example, it is interesting that two countries, Denmark and Finland, have demonstrated no RAEs for educational attainment. The authors (Bedard and Dhuey, 2006) note that in countries that demonstrated no RAEs, those countries also had a lack of, or in the case of Denmark a complete prohibition on, ability streaming until late adolescence (16 years of age). Furthermore, the results from my doctoral thesis (Wattie, 2011) suggested that school quality (as measured by Ofsted) may have an

---

5 The Office for Standards in Education, Children’s Services and Skills (Ofsted) is responsible for inspecting the care of young people, and the provision of education, skills for learners, and safeguarding of youth of all ages (www.ofsted.gov.uk).
influence on relative age’s influence on sport participation, with a low quality-rated school demonstrating RAES for sport participation among boys, but no indications of RAES within a high quality-rated school.

In summary, technical solutions might not address some of the causes of RAES. Rather than assume that solutions need to be technical, it may be constructive to consider human solutions, e.g. addressing beliefs in ability streaming, societal preoccupations with talent development/identification (in education and sport), inequalities in the distributions of wealth and resources, and a number of other institutions. However, relative age, like human development, does not operate in a univariate vacuum (Schulenberg, 2006). As such, future research will need to acknowledge that relative age cannot be extricated from wider social trends, and that there may be a need to move away from univariate de-contextualized research and consider other important socio-demographic variables alongside relative age.

The second issue is a general one, and concerns all social policies. The reality is that all social policies have the potential to produce unforeseeable negative consequences. RAES are a good example of an unanticipated consequence of annual age grouping policies. Therefore, it may be important to consider the unintended consequences of the technical solutions to RAES. Take the relative age quota solution: ensuring that an equal number of youth in each relative age Quartile are selected to a sports team. One of the realities of relative age is that considerable variability in physical maturation (height and weight) exists between youth of different relative age. There could be considerable injury risk for youth of different levels of physical maturation who participate together in a high speed collision sport. Similarly, rotating selection criteria (e.g., the Novem system or the RAF system) may adversely affect some of the very reasons why youth participate in sport in the first place - being with friends and having fun are the common reasons youths give (McCarthy et al., 2008, Weiss and Williams, 2004). The influence of repeatedly rotating selection-dates (thereby changing cohorts) on the social-motivational reasons for youth sport participation has yet to be explored, but could be an important consideration. Also, what effect would it have on youths self-perceptions when two children receive the same grade or mark on a test (age normalization of test results), but when performance is evidently different, as in PE (e.g., “Mark is faster than me”)?

Unfortunately the institutions which influence RAES may be complicated, and implementing solutions to RAES may be equally
complicated. As such, the prospect of redressing relative age inequalities seems daunting. However, where and for whom any institutions (i.e., annual age grouping policies, belief in ability streaming, and social inequalities) create RAES that are pervasive, enduring and of meaningful magnitude, there may be an ethical responsibility to address them. As Charles Darwin suggested, “If the misery of our poor be caused not by the laws of nature, but by our institutions, great is our sin” (Darwin, 1890, p. 596).^6

**References**


ABERNETHY, B. & FARROW, D. Contextual factors influencing the development of expertise in Australian athletes. 11th World Congress of Sport Psychology, Aug 15-19 2005 Sydney (NSW).


---

^6 At the risk of taking this quote out of context, it should be noted that Darwin was denouncing the far more serious and heinous social institution of slavery, which he witnessed during his voyage of the Beagle.


**Acknowledgements**

I would like to thank Dr. Jim McKenna and Dr. Steve Cobley (my Doctoral supervisors) for their support while completing my Ph.D. Furthermore, I would like to acknowledge the financial support of a Doctoral Fellowship from the Social Sciences and Humanities Research Council of Canada (SSHRC Grant # 752-2008-0280), a Sport
Participation Research Initiative Supplement from Sport Canada (Grant # 862-2008-0007), and a Centenary bursary from Leeds Met, that I received during my doctoral studies.

Nick Wattie, 111 Four Seasons Drive, Nepean, Ontario K2E 7R1, Canada

Email: wattie.n@gmail.com