The Relative Age Effect in Female Gymnastics: A Flip-Flop Phenomenon

Submitted to: *International Journal of Sport Psychology*

April 15, 2013

Resubmitted December 12, 2013

Second Resubmission February 18, 2014

David J. Hancock¹, Janet L. Starkes², Diane M. Ste-Marie³

1 Indiana University Kokomo; Kokomo, Indiana, USA

² McMaster University; Hamilton, Ontario, Canada

³ University of Ottawa; Ottawa, Ontario, Canada

Corresponding Author:
David J. Hancock
2300 South Washington Street, SM #106C
Indiana University Kokomo
Kokomo, IN, USA, 46904-9004
(765) 455-9234
dhancoc@iuk.edu

KEYWORDS: sport; female; talent development; birth date
Abstract

Relative age effects are pervasive throughout sport; however, little is known about how relative age interacts with other mechanisms of expert performance such as sports characterized by young ages for peak performance. The purpose of this study was to examine relative age effects in gymnastics, a sport where athletes reach peak performance during puberty. Examining the birthdates of 921 female gymnasts, we discovered no relative age effects for the collective sample. Dividing the sample into two groups, we noted a relative age effect for under-15 gymnasts, but a reverse effect for over-15 gymnasts. Inspecting competitive standards, we noted over-representations of over-15 athletes born in the fourth quartile at all standards except national. In the discussion, we highlight the complexity of relative age effects by incorporating deliberate practice and competition standard as variables for consideration.
One of the underlying goals of sport science research is to understand how to achieve sport expertise. In this vein, researchers have posited key requirements for attaining expertise including such features as deliberate practice (Baker, Bagats, Büsch, Strauss, & Schorer, 2012; Helsen, Starkes, & Hodges, 1998), deliberate play (Berry, Abernethy, & Côté, 2008), coaching (Baker & Horton, 2004; Côté & Fraser-Thomas, 2008), and a supportive family (Côté, 1999; Fraser-Thomas, Strachan, & Jeffery-Tosoni, 2013). Researchers have also proposed different pathways and trajectories that develop talent and lead to sport expertise including early diversification (Côté, 1999), early engagement (Ford, Ward, Hodges, & Williams, 2009), or early specialization (Law, Côté, & Ericsson, 2007; Ward, Hodges, Williams, & Starkes, 2004).

Among the many manuscripts related to sport expertise, a seminal paper by Starkes (2000) indicated that other factors beyond deliberate practice contribute to sport expertise. Specifically, Starkes suggested that while many coaches seek “talented” athletes who will excel under deliberate practice conditions, coaches’ notions of talent might be masked by a phenomenon known as the relative age effect. The relative age effect, which can be defined as the consequences of age variations between individuals competing in the same cohort (Musch & Grondin, 2001), arises as a function of the annual age-bands that many governing sport bodies set for youth sport participation (e.g., January 1 to December 31). What has been typically noted in elite sport is an over-representation of athletes born in the first three months following the cut-off date than others born closer to the end of the age-band, i.e., a greater proportion of relatively older athletes compete on elite sport teams as compared to their younger counterparts (Barnsley, Thompson, & Barnsley, 1985).

Relative age effects in sport were first reported in Canadian ice hockey (Barnsley et al., 1985; Grondin, Deshaies, & Nault, 1984). Since then, numerous studies have reported relative
age effects (see Cobley, Baker, Wattie, & McKenna, 2009 for a review). Despite a plethora of research, most studies focus on male team sports, often with the emphasis on replication. Studies across other sport contexts, however, are important to further illuminate and understand the complexity of this phenomenon.

As the field of relative age effects progressed, several authors (e.g., Helsen, Hodges, van Winckel, & Starkes, 2000; Schorer, Cobley, Büsch, Bräutigam, & Baker, 2009; Sherar, Baxter-Jones, Faulkner, & Russell, 2007) have agreed with Starkes’ (2000) assertion that coaches might mistake physical maturity (i.e., being relatively older) for talent, and that physical maturity is correlated with performance in youth. As such, it is important for sport scientists to further investigate relative age effects as they have substantial implications for the development of sport expertise. Thus, our purpose was to inspect the relative age effect in female gymnastics – an uncommon sample in relative age research.

Starkes (2000) highlighted that deliberate practice and relative age may interact and impact on expertise. Deliberate practice refers to effortful engagement in an activity that is not inherently enjoyable (Ericsson, Krampe, Tesch-Römer, 1993), but is ideal for accumulating relevant experiences to attain expertise. The time invested in deliberate practice differs dependent on the sport (Côté, Lidor, & Hackfort, 2009). In sports where peak performance is achieved during adulthood, it is recommended that most deliberate practice hours be acquired during late adolescence. In contrast, when peak performance is reached before adulthood, investing in deliberate practice during childhood and early adolescence is argued to be vital to success (Côté & Fraser-Thomas, 2008). Supporting this, Law and colleagues (2007) showed that Olympic gymnasts, whose peak performance is typically achieved at ages 15 to 17, engaged in more deliberate practice at younger ages than international gymnasts.
Given that deliberate practice is effortful, long-term, and not inherently enjoyable (Ericsson et al., 1993), it is reasonable to suspect that during childhood and early adolescence, relatively older athletes might excel more than relatively younger athletes. This is argued due to the increased cognitive maturity that may be needed for deliberate practice training conditions (Baker, Janning, Wong, Cobley, & Schorer, 2014). Yet in gymnastics, a smaller and lighter body can have distinct biomechanical advantages (for example, strength to body mass ratio) (Monsma & Malina, 2005), and therefore, relatively younger athletes might be better suited for certain types of performance. Such biomechanical advantages may also be more pronounced at the highest level of competition; this physical immaturity, however, would be competing against the possible advantages associated with the greater cognitive maturity requirements of deliberate practice that were previously suggested. These varied factors highlight the potential complexity of the relative age effect in gymnastics, and accentuate the need for continued study.

Considering these possible interacting factors in gymnastics, we hypothesized that relatively younger gymnasts would be advantaged (i.e., greater frequency of participation) at the lower levels in competition, but that at the highest competitive standards (where increased deliberate practice is required), relatively older gymnasts would be advantaged (i.e., over-represented on the most elite teams).

**Method**

**Participants**

Participants were 921 female gymnasts from Ontario, Canada. The data were collected in 1998 from the Ontario Gymnastics Federation and included the gymnasts’ birthdate and standard of competition (from highest to lowest: National, regional, provincial, and elite-provincial\(^1\)).

Before continuing, it is important to mention that although our data represent a sample of

\(^1\) An athlete could not be registered in multiple competitive standards.
gymnasts from 1998, there is still relevance in its publication. While other data on Canadian
gymnasts have since been published (e.g., Baker et al., 2014), our data are unique in that it
contains a very large, regional sample of athletes. Specifically, the sample in this study has over
900 athletes in one year competing at multiple competitive standards in one Canadian province,
compared to Baker and colleagues’ (2014) sample of less than 300 athletes over 10 years,
competing only at the national competitive standard – thus each study has distinct benefits. It is
also important to note that the Ontario Gymnastics Federation does not appear to have undergone
significant structural changes since 1998, thus we have no reason to believe that a more current
dataset would yield dissimilar results to those found herein. In fact, comparing our older
national-level sample to Baker and colleagues’ newer national-level sample, the same birth rate
distribution is evident.

The selection year for gymnastics in Canada runs from January 1 to December 31; thus,
birth quartiles (Q) were as follows: Q1 = January to March; Q2 = April to June; Q3 = July to
September; Q4 = October to December. To ease the interpretation of participants’ ages,
gymnasts were assigned an age depending on how old they would be at the end of the calendar
year. For instance, gymnasts born in 1988 were all classified as “10 year olds”. We further split
the participants into two distinct age groups. Group one was labelled “under-15”, while group
two was labelled “over-15”. This distinction was derived from Baxter-Jones and colleagues
(Baxter-Jones, Helms, Baines-Preece, & Preece, 1994), who noted that the average age of
menarche for female gymnasts was 14.3 years – approximately two years later than the average
population (Anderson, Dallal, & Must, 2003). Thus, our groups were, at least theoretically,
aligned with being pre-menarche and post-menarche.

Analysis
To test if the distribution of birth dates significantly deviated from the expected distribution, we implemented chi-square goodness of fit tests and calculated effect sizes ($w$) to determine the strength and meaningfulness of results. Cohen (1988) suggested that $w$ values of 0.1, 0.3, and 0.5 correspond to small, medium, and large effect sizes, respectively. Tests were conducted on (1) the overall sample, (2) an under-15 versus over-15 sample regardless of standard of competition, and (3) an under-15 versus over-15 sample across competitive standards. Often, researchers set the expected value of each quartile at 25%, or by using live birth rates, which have little variance (e.g., live birth rates in Canada range from 23.9% to 26.2% in each quartile; Statistics Canada, 2013). Another possibility is to derive expected quartile values from a known referent sample, such as comparing birth rate distributions of the most elite athletes to birth rate distributions of near-elite athletes (e.g., Schorer et al., 2009). In our study, we combined two methods for setting expected distributions. Specifically, for the overall sample, the under-15 total sample, and the under-15 regional sample, we set the expected quartile distributions at 25%, as the samples were likely drawn from the general population. For the under-15 provincial, elite-provincial, and national groups, the referent sample was the previous under-15 competitive standard (e.g., under-15 nationals compared to under-15 elite-provincials). Finally, for the over-15 groups, the referent sample was the corresponding under-15 group (e.g., over-15 nationals compared to under-15 nationals). This course of action was chosen as we aimed to investigate differences in relative age before and after age 15. In cases where the chi-square test was significant ($p < .05$) or approached significance, we employed post-hoc tests (standardized residuals) to identify which quartiles deviated from the expected distribution. This procedure was first used in relative age research by Hancock, Young, and Ste-Marie (2011) and provides a more robust interpretation of the results. Essentially, standardized
residuals act as z-scores, with scores greater than 1.96 indicating a significant \( p < .05 \) over-representation of athletes born in a particular quartile, and scores less than -1.96 indicating a significant \( p < .05 \) under-representation of athletes born in a quartile. Additionally, in cases of non-significant tests, we provided test power \( (1-\beta) \).

**Results**

A summary of the results is presented in Table 1. When testing the entire sample regardless of age and competitive standard, no relative age effect existed: \( \chi^2 (3, 920) = 3.42, p = .33, w = 0.06, 1-\beta = .70 \) (see Figure 1). The next test involved dividing the group into under-15 and over-15 gymnasts. The under-15 group demonstrated a significant relative age effect: \( \chi^2 (3, 735) = 9.21, p = .03, w = 0.11 \). Inspection of the standardized residuals revealed a relative age effect with athletes born in the second quartile being over-represented \( (SR = 1.99, p < .05) \) while athletes born in the fourth quartile were under-represented \( (SR = -2.05, p < .05) \). For the over-15 group, there again was a significant relative age effect: \( \chi^2 (3, 184) = 21.54, p < .001, w = 0.34 \); however, in this case, athletes born in the first quartile were under-represented \( (SR = -2.26, p < .05) \), while athletes born in the fourth quartile were over-represented \( (SR = 3.80, p < .001) \). In fact, 34% of over-15 registrants were born in the fourth quartile, which represented a 13% increase compared to the sample of under-15 gymnasts.

The final step was to test relative age for the under-15 and over-15 gymnasts based on competitive standard (see Figure 2). For under-15 regional athletes, no relative age effects were present \( (1-\beta = .72) \). When comparing the under-15 regional athletes to the over-15 regional athletes, however, there was a significant relative age effect \( \chi^2 (3, 73) = 9.47, p = .02, w = 0.36 \) with an over-representation of athletes born in the fourth quartile \( (SR = 2.37, p < .05) \). The identical trend existed at the provincial and elite-provincial standards whereby under-15 athletes
showed no relative age effects (1-β = .90 and .17 for provincial and elite provincial, respectively), while over-15 provincials (χ² (3, 61) = 8.44, p = .04, w = 0.37) and over-15 elite provincials (χ² (3, 27) = 9.03, p = .03, w = 0.57) each demonstrated significant effects with fourth-quartile over-representations (SR = 2.36 and 2.00, p < .05, respectively). Finally, we examined under-15 and over-15 national athletes, the highest competitive standard in our sample. At the national standard, neither the under-15 (1-β = .65) nor the over-15 (1-β = .68) group possessed significant relative age effects.

Discussion

The purpose of this paper was to examine the birth rate distribution of female gymnasts, thereby adding to the current literature on relative age effects. The two most interesting findings were: (1) a shifting relative age trend between athletes under 15 years of age (relatively older were advantaged) and athletes over 15 years of age (relatively younger were advantaged), and (2) a non-significant over-representation of national level athletes born in the second quartile. In this discussion section we integrate our results with previous relative age effect and expertise literature.

Our results align with previous studies on artistic sports, whereby it has been suggested that being relatively oldest is not automatically beneficial (e.g., Baker et al., 2014; van Rossum, 2006; Wattie et al., 2014). Specific to gymnastics, it has been shown that athletes born in the second and third quartiles were more frequently represented on national teams than athletes born in the first and fourth quartiles (Baker et al., 2014). Additionally, relatively older gymnasts have been shown to drop out more frequently than relatively younger gymnasts (Wattie et al., 2014). Therefore, it should not be surprising that our results showed advantages, at times, for relatively younger gymnasts. Wattie and colleagues (2014) suggested that puberty might play a role in
dropout, particularly psychological responses such as increases in depressive symptoms and weight concerns, as well as decreased feelings of self-worth. We also proposed that puberty might cause differences in relative age, and thus grouped our participants into under- and over-15 years of age. According to the literature on gymnasts, these groups would represent pre-pubescent and post-pubescent athletes respectively (Baxter-Jones et al., 1994). Specifically, the under-15 gymnasts possessed a relative age effect whereby those born earlier in the year appeared more frequently in the dataset, while over-15 gymnasts demonstrated a relative age effect as it favored those born later in the year. Again, this trend is unusual as most sports that exhibit relative age effects show fairly stable patterns throughout youth sport, though changes in the trend are sometimes witnessed in adult sport (see Ashworth & Heyndels, 2007; Wattie & Baker, 2013).

Considering this result, it is important to contemplate the distinct and unique aspects of gymnastics that create such a trend. One prevailing difference between gymnastics and other traditional team sports is the age at peak performance. It has been shown that in sports where elite levels of performance are attained prior to adulthood, there is an increased emphasis on deliberate practice (Côté et al., 2009). Specific to our population of gymnasts, Law and colleagues (2007) noted that they typically peak between ages 15 and 17, and thus engage in intense deliberate practice from very young ages in order to achieve elite performance standards. Alternatively, when expert performance occurs later in life, late twenties for instance, deliberate practice hours can be spread out over a longer timeframe. Indeed, Law and colleagues discovered that Olympic gymnasts trained 25 hours per week from ages 9 to 12 years, while those who competed internationally (but did not make the Olympics) trained for only 11 hours
per week from ages 9 to 12, which supports the idea that early amounts of deliberate practice are critical for success in gymnastics.

Related to our findings, it is possible that before puberty, relatively older athletes are more cognitively prepared for the rigors of deliberate practice than their relatively younger counterparts. Then after puberty, when there are smaller cognitive maturity discrepancies amongst competitors, deliberate practice can be engaged in intensively by all. As such, the post-puberty age advantage witnessed for relatively younger gymnasts may now be explained by the possible biomechanical advantages of the sport (Monsma & Malina, 2005). Specifically, athletes born in the second half of the year might be slightly less physically mature, resulting in shorter stature, “straighter lines” (leading to increased visual aesthetics), lower weight, and higher strength to body mass ratio than relatively older athletes. These physical characteristics would result in biomechanical advantages conducive for talent development and superior gymnastic performance (Monsma & Malina, 2005). Perhaps in female gymnastics there is an ideal level of physical maturity where athletes are capable of intense deliberate practice, but not so developed as to have their performances impacted.

A piece of the data that makes the picture less clear is the fact that, contrary to the lower levels of competition, national athletes – independent of age – showed no significant results, though there was an over-representation of athletes born in the second quartile. The lack of significance here was surprising, as was the consistent over-representation of second-quartile athletes. A plausible explanation also considers deliberate practice. Specifically, it is possible that national athletes who were born in the second quartile have an increased ability to engage in intensive deliberate practice (compared to third- and fourth-quartile athletes), perhaps resulting from greater cognitive maturity. Since they are not the most mature gymnasts (compared to first-
quartile athletes), they might also maintain some important biomechanical advantages. This
would create an ideal combination of ability (acquired through early deliberate practice and
coaching) and physical maturity, thereby preventing the shifting trend noted in the less elite
categories. This explanation is speculative and requires further attention. Certainly, more recent
explanations of relative age trends in artistic sports might also explain our results herein. As an
example, Baker and colleagues (2014) suggested relatively older female athletes might transfer
to other sports, while Wattie and colleagues (2014) postulated that psychological responses to
puberty might impact birth rate distributions. Clearly further attention to the interaction of age,
maturity, and practice in female gymnastics is warranted.

Finally, while our national results were not significant (which could be a sign of low test
power), the medium to large effects sizes make this an important contribution to a growing body
of research on female relative age effects, which once again highlight an over-representation of
athletes born in the second-quartile at elite levels of performance (see also Baker, Schorer,
Cobley, Bräutigam, & Büsch, 2009; Baker et al., 2014; Delorme, Boiché, & Raspaud, 2010;
Hancock, Seal, Young, Weir, & Ste-Marie, 2013; Weir, Smith, Paterson, & Horton, 2010). This
peculiarity has transcended sports, cultures, and countries, yet explanations (puberty,
psychological, sport transfer, etc…) are not conclusive. We recommend sport scientists dedicate
attention to understanding the second-quartile phenomenon.

While this study brings further understanding to female relative age effects in artistic
sport, there are limitations. First, in terms of the sample, the national level sample size is small,
making it difficult to draw firm, generalizable conclusions, and we recognize that more present-
day data could possibly lead to different outcomes. Second, we did not take any direct measures
of maturation; rather, we used previous literature to offer opinions on pre- and post-pubescent
athletes. Finally, we did not record deliberate practice hours, which will be required in a future study to truly understand the interaction of deliberate practice and relative age.

In conclusion, we have presented an examination of relative age effects in a sport context that sheds light on its complexities and impact on expertise. Factors such as competitive level, sport demands, and consequent intensity of deliberate practice were presented as important variables for consideration. The shifting relative age effects speak to the fact that sport scientists should continue this line of inquiry, perhaps taking more direct measures to test the interaction of relative age, physical and cognitive maturation levels, and deliberate practice.

References


practice. *Journal of Sport & Exercise Psychology, 20*, 12-34.


### Statistics for Chi-Square Tests

<table>
<thead>
<tr>
<th>Group</th>
<th>Chi-Square Statistics</th>
<th>Quartiles (Raw)</th>
<th>Standardized Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>$\chi^2$</td>
<td>$p$</td>
</tr>
<tr>
<td>Entire Sample</td>
<td>921</td>
<td>3.37</td>
<td>.33</td>
</tr>
<tr>
<td>Under-15</td>
<td>736</td>
<td>9.21</td>
<td>.03*</td>
</tr>
<tr>
<td>Over-15</td>
<td>185</td>
<td>21.54</td>
<td>.00*</td>
</tr>
<tr>
<td>Under-15 Regional</td>
<td>387</td>
<td>3.17</td>
<td>.37</td>
</tr>
<tr>
<td>Over-15 Regional</td>
<td>74</td>
<td>9.47</td>
<td>.02*</td>
</tr>
<tr>
<td>Under-15 Provincial</td>
<td>208</td>
<td>0.76</td>
<td>.86</td>
</tr>
<tr>
<td>Over-15 Provincial</td>
<td>62</td>
<td>8.44</td>
<td>.04*</td>
</tr>
<tr>
<td>Under-15 Elite-Provincial</td>
<td>85</td>
<td>5.58</td>
<td>.13</td>
</tr>
<tr>
<td>Over-15 Elite-Provincial</td>
<td>28</td>
<td>9.03</td>
<td>.03*</td>
</tr>
<tr>
<td>Under-15 National</td>
<td>56</td>
<td>5.69</td>
<td>.13</td>
</tr>
<tr>
<td>Over-15 National</td>
<td>21</td>
<td>4.62</td>
<td>.20</td>
</tr>
</tbody>
</table>

* Denotes significant at $p < .05$. ^ Denotes significant at $p < .01$. + Denotes significant at $p < .001$.

---

**Figure 1:** Relative age distribution for entire sample, under-15 athletes, and over-15 athletes
Figure 2: Under-15 and over-15 gymnasts according to competitive standard