Skip Counting Bases and Autonomous Motivation

Chapter submission for the proposed Nova volume

Goal Setting and Personal Development:
Teachers' Perspectives, Behavioral Strategies and Impact on Performance

Stephen E. Erfle
International Business and Management
Dickinson College
Carlisle, PA 17013

ABSTRACT

Histograms of push-up and curl-up performances of middle school students show that students stop at multiples of five more often than random processes suggest. The natural question arises: do students who stop at these focal outcomes settle for them or did they strive to achieve them? Erfle and Gelbaugh (2013) and Erfle (2014b) argue that the latter is true and that these students are more motivated than their less focal peers. This motivation to perform transfers across activities, especially for focal push-up performers. The present chapter examines whether students implicitly use other counting bases in their physical activity performances. The relative performance of those who count by other bases are also examined.

A performance ending in a multiple of k may not be the result of an intentional strategy of skip counting by ks. In the absence of student survey information regarding skip counting strategies we must infer intentionality to perform using a specific base from actual performance outcomes. As a result, one task is to define intentionality to attain or avoid a specific outcome. If N students complete performances and k is a counting base, then having more than N/k students who end their performance in multiples of k provides prima fascia evidence of counting by k. Similarly, having fewer than N/k students who end their performance just below multiples of k offers further evidence of counting by k. Evidence of both behaviors exists for a number of bases for both activities. Average performance profiles for those who end in multiples of k as well as those achieving a remainder of k-1 are examined as well.

INTRODUCTION

This chapter lays out a methodology for analyzing skip counting by various bases. But this methodology was only conceived of because initial analysis of the United States Service Academy at West Point Cadet data analyzed in the next chapter suggested that, unlike middle-school students analyzed by Erfle and Gelbaugh (2013) and Erfle (2014b), Cadets exhibit little proclivity for counting by

5s. Base 5 and 10 were the obvious choices given the focal spikes at multiples of 5 and 10 in the middle-school data analyzed in Erfle (2014b). These performances are summarized by the histograms in Figure 1.

******* Figure 1 about here *******

Some physical activities require that performances end in a whole number, n. This outcome can always be achieved as a result of counting by ones n times or counting by n one time. If n is a composite number, then this outcome can be achieved in multiple ways via skip counting by various bases.

Skip counting is counting by a number larger than one. This counting technique is taught as a way to speed up manual counting, increase accuracy when counting a large numbers of objects, and as a way to learn multiplication (Campbell & Graham, 1985). A performance of 12 can be achieved as a result of skip counting by 2s, 3s, 4s, 6s or 12s and 20 can result from skip counting by 2s, 4s, 5s, 10s or 20s. Both of these outcomes could, of course, also be attained simply as the result of counting by ones. This chapter examines whether credible evidence exists that suggests that students are using skip counting as they undertake physical activity performances ending in whole number outcomes.

Erfle and Gelbaugh (2013) and Erfle (2014b) showed that students who used focal counting on one event (remainder 0 upon division by 5 for push-ups; remainder 0 upon division by 10 for curl-ups) were more likely to do so on another event, or on the same event at a later assessment. They also found that students who stopped at these focal endings outperformed their non-focal peers on fitness tests. They found that males were more likely to be focal than females and that focal proclivity increases with age.

Because motivation is unobservable, it must be inferred from indirect measures. Some studies have relied on questionnaires to infer the nature of motivation while others have suggested behavioral metrics for autonomous motivation (Katz, Assor, & Kanat-Maymon, 2008; Mayer, Faber, & Xu, 2007). Erfle and Gelbaugh (2013) argue that systematic focal behavior acts as one such behavioral metric for autonomous motivation.

METHOD

Participants

Although this analysis was spurred by the lack of systematic behavior at multiples of 5 in the Cadet data analyzed in the next chapter, the present analysis is performed on a dataset that was created over a period of two academic years to analyze the question: does 30 minutes of daily physical education, PE, impact the fitness and obesity status of middle-school students relative to those do not have daily PE? Erfle (2014a) and Erfle and Gamble (2015) provide answers to that question. Those articles lay out in detail the attributes of the Pennsylvania Department of Health/Active Living Research *Active Schools Program*, ASP, dataset. The basic attributes of that data is described below.

The Pennsylvania Department of Health instituted the ASP to encourage daily physical activity among middle-school students. Participating schools agreed to institute a regimen of at least 30 minutes of daily PE and to assess physical fitness and weight status at the start and end of the academic year. The Robert Wood Johnson Foundation funded a control school analysis of the ASP program through Active Living Research Rapid Response Grant #68311. Control schools administered the ASP assessment protocol at the start and end of the academic year but otherwise maintained their regimen of non-daily PE. The combined ASP and control schools dataset had 10,206 students from 39 schools with complete fall and spring information (the curl-up and push-up performances of these individuals is shown in Figure 1). The present analysis is restricted to the 9,345 students who performed at least one curl-up and one push-up for the fall and spring assessments.

Procedure

Excel was used for data cleaning and SPSS was used for statistical tests. Tests include chi-square tests, independent sample *t* tests and one-way analysis of variance tests paired with the LSD method of multiple comparisons. This chapter uses $p < 0.05$ to test for statistical significance.

The first task is to consider how to describe specific systematic behavior when that behavior can be the result of multiple skip counting options as well as counting by ones.

Proclivity revisited. Erfle and Gelbaugh (2013) distinguished two aspects of focal performance: proclivity and performance. Proclivity was interested in assessing the issue of whether random processes could have led to the outcome under consideration. Specifically, do push-up performances end in multiples of 5 and curl-up (also called sit-ups) performances end in multiples of 10 more often than random processes would suggest. Performance examined whether those who stopped at these focal outcomes performed differently from those who did not on various events.

The spikes in the histograms from Figure 1 provide visual proof that students count by 5s more often than random processes would suggest but formal tests of this proposition confirm the visual image. More importantly, tests confirm that being focal on one event makes it more likely that the individual would be focal on the other event. Erfle and Gelbaugh (2013) showed that those who were focal on both events ran the mile about fifteen seconds faster than those who were not. This result was significant at the 5% level. By contrast, being focal on one event was not a significant positive predictor of mile performance. They concluded that systematic behavior was required for benefits that move across activities. Erfle (2014b) added a second assessment thereby doubling the number of events examined and increasing by a factor of four the number of possible methods of achieving systematic focal behavior. He showed that systematic focal push-up performances are more indicative of cross-activity superior performance than systematic focal curl-up performances.

Once we consider multiple bases we need to create a methodology for comparing across bases. One method to compare across bases is to use standardized residuals. A residual, r, is the actual number of members of a partition bin, g, minus its expected value, E, or r = g – E. A standardized residual, s, is used to standardize these deviations (i.e. residuals) across different sized expected values. It is defined as:

$$s = r/E^{0.5} \qquad (\mathbf{1})$$

In statistics, the sum of squared standardized residuals equals chi-square. A chi-square test allows one to examine whether there are systematic differences across cells in a partition of data relative to expectations. This test statistic treats positive and negative residuals symmetrically. We now propose a measure which treats positive and negative residuals asymmetrically in order to have a symmetric measure of how intentional are the residual values. This is the notion of intentionality to which we now turn.

Defining Intentionality. We can ask the question, how much intentionality can be seen with regard to counting by base k in a given sample of N students. There are k possible remainder groups when a score is divided by k. Each student will populate a remainder group even in the absence of skip counting tendencies.

A simple example will help clarify the situation. Suppose no students have a counting strategy. In this event, we would expect half of the students to stop at even numbers (remainder group 0) and half to stop at odd numbers (remainder group 1), at least that should be the result for large samples. A chi-square test readily confirms whether significant differences exist between remainder groups but we are left to wonder how much of the difference is due to intentional behavior and how much is due to random behavior. If 50% of students ended at even outcomes then we could infer that no intentionality was involved because this is what we would expect to happen under the assumption of random processes. Conversely, if all students ended at even outcomes then we could infer full, or 100% intentionality. Between these bounds we can infer partial intentionality.

Scenario I: Suppose 70% of the students end at even outcomes.

It would be inappropriate to say that 70% of students planned to stop at even outcomes because 50% would be expected to end there as a result of random processes. A more appropriate view is to say that 50% are due to random processes and the remaining 20% are intentional in nature. This would imply that 40% of all students were intentional, not 20%. The reasoning is straightforward, assume that all 40% of students who counted by twos ended up at an even outcome. Half of the remaining 60% would also end up at even outcomes as a result of random processes, yielding 70% of all students ending at even outcomes in this instance (70% = 40% + 60%/2).

More generally, if we consider counting by base k with N students, then we would expect E = N/k students in each of the k remainder group bins from i = 0, 1, …, k-1 based on random processes. If we let g be the actual number of students in a bin the residual for that bin is r = g – E.

For $r \geq 0$, intentionality to attain this outcome is: $I = r/(N - E)$. (**2**)
For $r \leq 0$, intentionality to avoid this outcome is: $I = r/E$.

Note that if $g = N$ then $I = 100\%$ and if $g = E$ then $I = 0\%$. If $g = 0$ then complete avoidance is attained, $I = -100\%$, and if $g = E$ there is no avoidance, $I = 0$. Put another way, intentionality ranges from -100% to 100% as g ranges from 0 to N. Unless each of the k remainder groups have exactly the expected outcome, then there will be at least one positive remainder group and one negative remainder group. There will also be a positive chi-square value which is evaluated with k-1 degrees of freedom to see if significant differences exist across remainder bins.

Our initial interest is in how intentional are remainder zero outcomes because this represents clear counting by k. If the zero remainder group has a positive bin residual then of secondary interest is whether there are systematic differences in the other k-1 non-zero remainder groups, 1 to k-1. For example, if just above (remainder group 1), or just below (remainder group k-1) is systematically more or less likely to occur than expected on the basis of random processes then this provides further evidence of systematic counting by base k. Consider two scenarios with 1,000 students and skip counting by k = 5.

Scenario II: $g_0 = 600$ and 100 are in each of the remaining four remainder groups.

This represents 50% intentionality to attain 0 (.5 = 400/[1000-200]) and 50% intentionality to avoid each of the other remainder groups (-.5 = -100/200).

Scenario III: $g_0 = 600$, $g_1 = g_4 = 0$, and $g_2 = g_3 = 200$.

Now there is 100% intentionality to avoid just above and just below outcomes, no intentionality with regard to the two “middle” values 2 and 3, as well as 50% intentionality to attain multiples of 5. Certainly Scenario III exhibits greater evidence of intentionality to count in 5s than was exhibited in Scenario II but it would be hard to place a single number on this result.

It is worth noting that a performance of zero on an event would be counted in the zero remainder group for all counting bases. Including performances of zero would bias the size of the zero remainder group. In order to avoid this bias, as noted above, the analysis is restricted to the 9,345 students who performed at least one repetition on each event. A similar strategy is followed in the next chapter with the Cadet data.

Restricted intentionality. A second way to analyze the intentionality of the nonzero remainder groups when the zero remainder group exhibits intentionality is to redo the intentionality analysis restricted to the nonzero remainder groups. Let g_0 be the actual count of students who are focal with regard to k and suppose $g_0 > E$ so that there is some intentionality to count by ks. The remaining $N – g_0$ students are not focal with regard to k. Given k – 1 nonzero remainder groups, the expected outcome in a nonzero groups based on random processes across nonzero remainder groups is $E_{nz} = (N – g_0)/(k-1)$ and the restricted nonzero remainder is $r_{nz} = g – E_{nz}$. Restricted intentionality, I_R, is:

For $r_{nz} \geq 0$, restricted intentionality to attain an outcome is:

$$I_R = r_{nz}/(N – g_0 – E_{nz}).$$

(3)

For $r_{nz} \leq 0$, restricted intentionality to avoid an outcome is:

$$I_R = r_{nz}/E_{nz}.$$

Consider once again Scenarios II and III using restricted intentionality. In Scenario II, I_R = 0% for all four remainder groups because each equals the nonzero expected value, E_{nz} = (1000 – 600)/4 = 100. I_R = 0 using both equations because the restricted remainder is zero when g = E_{nz}. In Scenario III, groups 1 and 4 have 100% restricted intentionality to avoid but the restricted intentionality to achieve for groups 2 and 3 is I_R = (200 – 100)/(1000 – 600 – 100) = 33% rather than 0% using the unrestricted definition of intentionality in Equation 2. It is worth emphasizing that restricted intentionality is conditional on g_0.

Performance profiles of skip counters. One can test for proclivity of counting bases for push-ups in bases beyond 5 but the level of difficulty of performing this particular activity is such that larger bases appear unreasonable from a practical perspective. The mean push-up performance for the 9,345 students doing at least one repetition on each event was 15.2 in the fall and 18.4 in the spring (with medians of 13 and 16). Focal performers with bases larger than 5 would have at most two focal opportunities prior to reaching median performance. By contrast, mean curl-up performance was 36.1 in the fall and 40.8 in the spring (with medians of 35 and 40) meaning that bases as large as 11 allow at least three focal opportunities prior to reaching median performance.

For certain bases, the fact that we have focal counting in excess of random processes for a given base may not provide useful performance information because the base is simply too large relative to the size of average performance. For example, 770 students in the fall and 831 students in the spring end their push-up performance in multiples of 15, 147 and 208 more than the 623 = 9345/15 expected on the basis of random processes (these focal outcomes have standardized residuals of s = 5.9 in the fall and s = 8.3 in the spring making base 15 the third largest remainder group). It is not surprising to find that the mean push-up performance for these students exceeds the mean push-up performance by those whose fall push-up performance did not end in a multiple of 15 ($\text{mean}_{\text{FallFocal,15}}$ = 22.8, $\text{mean}_{\text{FallnotFocal,15}}$ = 14.6, and $\text{mean}_{\text{SpringFocal,15}}$ = 24.4, $\text{mean}_{\text{SpringnotFocal,15}}$ = 17.8). This result is tautological in the fall because the counting base exceeds the mean performance of the rest of the students. No bases other between 5 and 15 other than 10 have proclivity levels exceeding that obtained from random processes for push-ups.

RESULTS

Proclivity of various bases

Figure 2 describes proclivity results for push-ups and curl-ups for both the fall and spring assessment. The upper four panels show standardized residuals for all counting bases k from 2 through 10 and the lower panel summarizes this information on a single intentionality graph using the above definitions. As noted in the previous section, one should consider skip counting push-up performances by bases above 5 with caution due to mean performance levels for push-ups. For no counting base between 5 and 10 is the remainder zero residual positive for push-ups. For curl-ups, 8 exhibits positive proclivity and both activities exhibit intentionality for counting by 2 and 4 due to positive residuals for remainder zero for those bases.

Intentionality. Unsurprisingly, bases 5 and 10 exhibit the greatest intentionality to achieve a focal outcome with intentionality values between 13% and 17% for base 5 and 8% to 11% for base 10. The only other base exhibiting focal intentionality in this range is base 2 where intentionality values range from 5% to 9%.

******* Figure 2 about here *******

Equally important is the pattern of negative residuals for the just below remainder groups. Especially impressive are the magnitudes of negative remainders in the base 5 and base 10 situations. In the base 10 scenario, there is a positive residual for remainder group 5 (i.e., ending at 5, 15, 25, etc.) but negative residuals for remainder groups 4 and 6 through 9 for all four events. The residual gets more and more negative as the remainder group approaches 9. Intentionality to avoid 9 is 30% to 35% for curl-ups and 45% to 50% for push-ups. In both instances, this avoidance is greater in the spring than the fall. In general, intentionality to avoid is greater than internationality to attain. This is not surprising when one considers the greater magnitude that a positive residual must be relative to a negative residual to exhibit the same percentage intentionality, especially when one is considering large counting bases.

Figure 3 compares focal (remainder zero) with just below (remainder k-1) by bases from 2 to 11. This is viewed from three different perspectives in Figure 3: (a) standardized residuals, (b) contribution to chi-square, and (c) intentionality. Focal base 3, 7 and 9 for push-ups have among the highest contributions to chi-square from the focal remainder group (except for 5 and 10) and each of these counting bases has a negative standardized residual for this outcome. Just below avoidance is greatest for counting by bases 5 and 10. Finally, note that for each event, the focal residual for base 11 is negative.

******* Figure 3 about here *******

Restricted Intentionality. As Figures 1 through 3 show, counting by 5s and 10s are the two most important counting bases for this data. Figure 4 provides an analysis of restricted intentionality for these counting bases. All three columns remove remainder 0 outcomes and examine how the nonzero remainder outcomes are distributed relative to one another. The left column of figures depict restricted intentionality in counting by 5s. The middle and right columns depict counting by 10s. Because counting by 10s has, as a half-way point, counting by 5s this is viewed in two methods in Figure 4, the middle panels exclude remainder 0 but include remainder 5 and the right column panels exclude both remainders 0 and 5.

******* Figure 4 about here *******

In each setting, the early remainder outcomes have greater restricted intentionality than the larger remainder outcomes. Each of the 12 chi-square tests (4 Event $\times$ 3 versions of nonzero remainder) are significant at the $p < .001$ level – there are significant systematic differences across nonzero remainder groups. In the middle panels we note that 30% to 66% of the chi-square contribution is based on the remainder 5 group. Clearly, some substantive fraction of individuals count by 5s and stop at odd multiples of 5 and restricted intentionality to attain this remainder stands at 5% to 7% depending on event in the middle panel of Figure 4, Part c. Once remainder 5 is removed and we focus on remainders 1-4 and 6-9 in the right column panels we see that the only remainders that have greater than average chi-square contribution are 1, 2, and 9 for push-ups and 2 and 9 for curl-ups. In each instance, 1 and 2 exhibit restricted intentionality to attain and 9 exhibits restricted intentionality to avoid. Avoidance of remainder 9 stands at 18% to 21% for curl-ups and 37% to 39% for push-ups. Put simply, many people who are at a last digit of 9 apparently push themselves to do an extra repetition in order to end in a multiple of 10 rather than end in 9, 19, 29, etc.

Skip counting by 5 and 10. The tendency to end in multiples of 5 and 10 can be thought of as the outcome of skip counting by fives. Ten turns out to a more important outcome than five, a result that can be seen using cross-tabs shown in Table 1. Most basically, this table is organized with an upper half and a lower half. The upper half examines even/odd comparisons and the lower half deals with bottom half/top half (remainder 0-4 versus 5-9 when divided by 10). Shaded portions in each relate the two halves to one another (so that the shaded cells in the upper half are the left column in the lower half and visa-versa). Cell counts and standardized residuals for a number of comparisons are provided within the table so it is worth pointing out the various tests that are considered within Table 1. There are a total of 40 chi-square tests reported in Table 1, 10 for each of the four events. Half of the tests include remainders 0 and 5 and half exclude remainders 0 and 5 upon division by 10. The former tests are in the five rows starting with **0-4 Total** and the latter are in the five rows starting with **1-4 Total**. Statistically significant chi-square tests and their associated standardized residuals are shown in bold. All told, 34 of the 40 tests are statistically significant, 29 of them at the $p < .001$ level.

******* Table 1 about here *******

By comparing chi-square tests with one another, we can see which cross-tabs exhibit the greatest deviation from independence across cells. Although each of the 1×4 chi-square values are significant at the $p < .001$ level, they are substantially lower than the 1×5 chi-square values. This suggest that much of the systematic difference across base 5 remainder cells is due to the remainder zero cell. This is confirmed in the even/odd 1×2 chi-square comparisons where all four are significant a $p < .001$ with 0 and 5 included, but only one of four is significant (fall curl-ups) once these values are excluded. The 1×2 chi-square comparisons in the bottom portion show that push-ups have a greater bottom/top effect than curl-ups, as expected given the difficulty of push-ups relative to curl-ups when viewed through a base 10 lens. Removing remainders 0 and 5 causes a larger change in standardized residual value for curl-ups than push-ups. When comparing 1×2 chi-square between even/odd versus bottom/top, we see a greater divergence (larger chi square value) for bottom/top than even/odd although that difference is more dramatic for push-ups than curl-ups. By contrast, when comparing 2×5 and 2×4 chi-square between even/odd versus bottom/top, we see a greater divergence for even/odd than bottom/top although that difference is more dramatic for push-ups than curl-ups.

In each of the four even/odd 2×5 cross-tabs the (0, 0) cell has a large positive residual relative to the (1, 0) cell. This is true even for push-ups despite the overall trend in push-ups that is skewed towards small performance outcomes. The same is true in the bottom/top 2 × 5 cross-tabs for both curl-up performances but not both push-up performances. The lone 2 × 5 chi-square that fails to achieve significance is push-ups in the spring using bottom/top as the "2" part of 2 × 5. This says that once one controls for bottom/top considerations, there are no significant differences across base 5 remainder groups (although with p = .053 one could call this result marginally significant). We see this same result in the restricted intentionality patterns in Figure 4. It is worth noting, however, that the 1,825 in the (0-4, 0) cell is less than expected using this categorization scheme (s = -0.3) but each of these students achieved at least 10 push-ups in order to be in that cell. By the same token, using an even/odd categorization, these 1,825 members of the (0, 0) cell are more than expected (s = 3.7) for that cell (because the even values in the 0 column of the 2 × 5 spring push-up even/odd cross-tab included last digit 6 and 8 rather than the more populated bottom values in the 0 column of the 2 × 5 spring push-up bottom/top cross-tab of 1 and 3).

Performance across various bases

Even in the absence of proclivity for counting by a specific base, we can examine whether performance varies across remainder groups for the Own-activity × Time as well as across other activities or times. Erfle (2014b) showed that base 5 focal push-up performances were associated with superior performance across events to a greater extent than were base 10 focal curl-up performances. For the reason discussed above, we restrict our attention here to bases 2-5 and 10 for push-ups.

Performance by all remainder groups. Figure 5 presents performance by remainder group in three panels, one for each physical activity. Mean value with 95% confidence interval whiskers for each remainder group are shown in four rows, the top two show fall performance and the bottom two show spring performance. The first and third rows describe fall focal remainder groups and the second and fourth show spring remainder groups. Own-event scenarios are shaded in the first two panels (the second half of rows 1 and 4 in the curl-up panel show fall and spring curl-up remainders, respectively but the first half of rows 1 and 4 in the push-up panel show fall and spring push-up remainders, respectively).

******* Figure 5 about here *******

The performance difference uncovered for base 5 push-ups in Erfle (2014b) relative to curl-ups base 10 appears to hold for other bases as well. Consider, for example, the effect of being focal base 4 for push-ups versus curl-ups on mile performance. In each row, the mean mile performance in remainder group 0 is lowest and the highest is remainder group 1 with groups 2 and 3 being stepped down in between these bounds. By contrast, no substantive systematic pattern emerges for base 4 focal curl-up remainder groups with regard to mile run performance. Another interesting general trend in Figure 5 is that focal performance appears to have the highest mean performance for most bases, but a very common next highest mean is the Just Below remainder group. As shown in Figure 2, this remainder group exhibits the greatest intentionality to avoid. But those who end in remainder k-1 appear to have superior performance to other groups except the focal remainder group 0.

Performance by the Focal, Middle, and Just Below partition. Rather than examine the difference between means for all remainder groups for each skip counting base we simplify the analysis by partitioning the data into a three remainder group partition: the Focal bin comprised of students having remainder 0 when the event performance is divided by k, the Just Below bin has remainder k-1 when the event performance is divided by k and the Middle bin (for k > 2) has remainders 1 to k-2 when the event performance is divided by k. The proclivity of the first two of these remainders was examined above in Figure 3. Figure 6 examines the relative performance on each activity using this three-way partition for the same bases covered in Figure 5. Three panels are shown, one for each of the two-way difference between means test comparisons. The three comparisons are (a) Focal – Just Below, (b) Focal – Middle, and (c) Just Below – Middle in Panels 6a through 6c, respectively. In each panel, a separate graph is shown for each activity with own-event focal performance differences shaded in Panels 6a and 6b; all other portions of each panel examine cross-event mean differences.

******* Figure 6 about here *******

Systematic differences emerge across the three-remainder partition. Because the general view from Figure 5 is that Focal performers have the highest mean performance followed by Just Below performers, the pair-wise difference between means comparisons in Figure 6 depict the expected dominant performer as the minuend (number that is to be subtracted from) in each panel (i.e., Focal is minuend in Panels 6a and 6b and Just Below is minuend in panel 6c). We see superior performance (positive mean differences on curl-ups and push-ups and negative mean difference on mile run [center of the difference between means 95% CI whisker]) for 72% of the 168 comparisons in Figure 6a, 74% of the 144 comparisons in Figure 6b, and 58% of the 144 comparisons in Figure 6c. Many of these differences are significant (i.e. the entire whisker is on one side of 0 or the other). Given the large number of comparisons shown in this figure, the results are summarized in tabular fashion in Table 2.

Table 2 is organized by focal performance with push-ups in the top half then curl-ups in the bottom half of the table. Within each half, rows are organized by activity (curl-ups, push-ups, mile run). Own-event performances are shown in gray-scale in Figure 6 and Table 2 and are separated from cross event performances for the same activity. The three panels of Figure 6 are shown as the three vertical thirds of Table 2. Each third is comprised of three columns which show the percent of all comparisons exhibiting this differential performance pattern. First, when a performance whisker is completely above the horizontal axis for push-ups and curl-ups or below the horizontal axis for mile run, then the minuend dominates the subtrahend. Second, when the reverse holds true, the subtrahend dominates the minuend. Third, when the whisker crosses the horizontal axis, no significant difference is observed. The number of comparisons for Figure 6a are shown to the right of the first three columns in Table 2 and between the second and third thirds for Figure 6b and 6c.

To clarify the organization, consider the 25% of comparisons in Figure 6a that show no significant difference for curl-ups given focal push-ups (the third entry in the first third of the first row in Table 2). These five comparisons (5 = 25% of 20) are the ones in the Focal push-ups portions of the curl-ups graph of Figure 6a that cross the horizontal axis. These five comparisons are: (a) focal push-ups in the fall base 4 for fall curl-ups, (b and c) focal push-ups in the fall bases 3 and 4 for spring curl-ups, and (d and e) focal push-ups in the spring bases 4 and 10 for fall curl-ups. The remaining 15 curl-up performance comparisons (15 = 75% of 20) of Focal versus Just Below show that Focal push-up performers dominate Just Below push-up performers on curl-ups (each of these 15 comparisons have whiskers that remain above 0).

******* Table 2 about here *******

Figure 6 and Table 2 show the asymmetry that exists between skip counting on push-ups versus curl-ups. Note that for none of the focal push-up comparisons does the subtrahend dominate the minuend across the three panels (i.e., the perverse result examined in the middle column of each third). By contrast for focal curl-up comparisons, Just Below dominates Focal 5.6% of the time across activities in Figure 6a, Middle dominates Focal 10.4% of the time across activities in Figure 6b, and Middle dominates Just Below 16.7% of the time across activities in Figure 6c. Focal push-ups comparisons on push-up performance exhibit significant comparisons every time (20 out of 20 comparisons in Figure 6a and 16 of 16 comparisons in Figure 6b). In Figure 6c, 5 of 8 own-event and 3 of 8 cross-event push-up comparisons show dominance by Just Below relative to Middle. Symmetric significant differences are much less common for those with focal curl-up performances. Own-event curl-up performances show dominance by the minuend 38.9%, 43.8%, and 37.5% of comparisons in Figures 6a, 6b, and 6c, respectively. Additionally, these smaller dominant performances are balanced by subtrahend dominant performances of 16.7%, 18.8%, and 37.5% of comparisons in Figures 6a, 6b, and 6c, respectively.

One may argue that the mile run provides the best indicator of overall benefit of focal counting because this activity is not related to either activity used to define focal counting. The mile run is necessarily a cross-event to both curl-ups and push-ups. Nonetheless, in each panel of Figure 6, the percent of comparisons with minuend dominance is greater for focal push-ups than curl-ups (65% versus 13.9%, 93.8% versus 9.4%, and 50% versus 6.3% in Figures 6a, 6b, and 6c, respectively).

Comparing Prime and Composite Outcomes

One of the difficulties with the above analysis is that performances ending in a composite number may be achieved as a result of skip counting using multiple bases, as well as simply counting by ones. The same cannot be said if the performance ends in a prime number.

We now turn to an analysis of the proclivity and relative performance of prime versus not prime performance outcomes. Each of the four events are examined separately in this analysis. Various subsampling methods are used to examine this topic from multiple perspectives. For each subsample, proclivity is measured relative to the number of possible outcomes of each type in the range, not relative to the actual number of outcomes of each type achieved under each option. For example, the prime number 113 and the composite number 116 are both *possible* outcomes in the sense that both numbers are below the maximum performance for each of the four events. Although a possible outcome for each event, neither number was an *actual* outcome in any event. Rather than exclude outcomes that did not occur for a given event, all possible outcomes were considered when determining whether actual outcomes are distributed randomly between prime and not prime values within the range for a given sample (not prime numbers are composites together with 0 and 1). Consider, for example, the Full sample fall curl-up performances. There are 202 possible outcomes between 0 and 201, 46 of which are prime and the remaining 156 are not prime. Students actually achieved 106 of those performance outcomes 25 of which were prime and the remaining 81 were not prime.

The first subsampling technique examines the difference between counting possible outcomes versus actual outcomes by removing high performance outliers. Specifically, the upper bound is the largest value having at least one student perform at this level and at all prior levels (so that the upper bound for fall curl-ups is 76 because no student performed 77 curl-ups in the fall and at least one student performed all outcomes prior to 77). This subsampling technique has no distinction between possible and actual prime versus not prime outcomes because all outcomes occurred in this unbroken series. This technique reduced the sample by less than 1% in each instance. A second subsampling technique, Full not 0, uses the 9,345 students achieving at least one repetition on all four events. This is the only subsample in Table 3 that is common across events and is the same subsample examined in Tables 1 and 2 and Figures 2-6 above.

Figure 2 and Table 1 have provided evidence of significant proclivity of counting by 2, 5 and 10. As a result, the next subsampling method includes only performances ending in odd numbers other than 5 (i.e., 1, 3, 7, or 9). This reduces the sample by 69% to 70% with the remaining performances being either an odd prime other than 5, or odd multiples of odd bases, or 1. A fourth subsample examines performance by students in the top half of each event by using median performance as the lower bound. Two subsamples exclude outliers on both ends of the performance spectrum by approximately equal percentages (of 10% and 20%). The final subsampling technique uses the second composite number, 6, as the lower bound (to avoid the first counts of 2, 3, 4, and 5) and adjusts the upper bound of included performances by explicitly using the largest prime performance as the upper bound. This subsampling technique excludes approximately 1% of curl-up performances and 25% of push-up performances. The reason for this difference is straightforward: 27.6% of fall push-up performances and 21.8% of spring push-up performances are less than six but 1.2% of fall curl-up performances and 0.8% of spring curl-up performances are less than six. Table 3 provides the proclivity and relative performance results of not prime relative to prime performance outcomes. The first half of the table examines fall performances and the second half examines spring performances.

******* Table 3 about here *******

Table 3 examines 32 two-way comparisons, eight for each of 2 Activities × 2 Assessment times. If we expect students to use skip-counting by various bases, then performances ending in a prime number should occur less often than those ending in not prime outcomes. Put another way, we should expect the prime residual to be negative and prime solutions will exhibit intentionality to avoid. This is true for 13 of 16 curl-up subsamples, 12 of which are statistically significant according to a chi-square test. Two of the three subsamples with perverse curl-up proclivity (shaded in grayscale) are the Odd not 5s subsamples. Both exhibit significant differences. Note further that the Odd not 5s subsamples are the only ones where prime possibilities exceed not prime possibilities. By contrast, less than half of the push-up subsamples exhibit the expected proclivity results (five of which are statistically significant). Only three of the subsamples exhibit proclivity in the expected direction for both assessment times: the Median or above subsample, the 20% Trimmed subsample, and the Largest prime subsample. Across assessment times and activities, the largest intentionality to avoid a prime outcome is obtained by curl-ups in the Unbroken series subsample and the largest intentionality to attain a prime outcome is push-ups in the Odd not 5 subsample.

To the extent that skip counting acts as a behavioral metric for intrinsic motivation and increased performance, we would expect performance differentials to tilt in the not prime direction. This should occur even in the absence of evidence of intentionality to avoid prime performance outcomes. We would expect positive mean differences in the third from bottom row of both panels of Table 3 given the mean difference is defined as:

$$\Delta \text{Mean} = \text{Mean}_{\text{NotPrime}} - \text{Mean}_{\text{Prime}}. \qquad \textbf{(4)}$$

This mean difference is positive for 12 of 16 subsamples in the fall assessment and for all 16 subsamples in the spring assessment. When we look across events for a given time period, the push-up differential is larger than the curl-up differential for 7 out of 8 comparisons for each assessment. This happens despite mean performance levels for push-ups being less than half the size of curl-up performance levels in all but the Median or above and the Largest prime subsamples. When we reverse course and look at differential performance across time for a given subsample and activity, we see that spring performance differential is larger than fall differential for 7 out of 8 comparisons for each activity.

CONCLUSION

With the definition of intentionality proposed in Equation 2 above, we see strong support (intentionality greater than 10%) for focal counting by 5 and 10 and weak support (5% to 10%) for counting by 2. Students end at multiples of 3, 7 and 9 less often than random processes would suggest for both activities (meaning that the residual is negative) and students end in multiples of 6 and 8 more often than random processes suggest for curl-ups but less often for push-ups. In none of these instances does intentionality to achieve the focal outcome attain 5%.

The present chapter provides further evidence in support of the assertion by Erfle and Gelbaugh (2013) that focal counting acts as a behavioral metric for intrinsic motivation. Focal performers using various bases for focal counting do, in general, out-perform their non-focal peers on own-event performances. On cross-event performances, focal push-up performers outperform their non-focal peers on other activities but the same cannot be said of focal curl-up performers. This is consistent with the activity asymmetry observed by Erfle (2014b).

One striking result from the intentionality analysis is the degree to which students avoid just below remainder outcomes. This avoidance is common across even counting bases from 2 to 10 as well as 5 and it appears to be the most strongly avoided outcome in each instance. The only bases where the just below outcome has a positive residual is base 3 (for all events), base 7 for spring push-ups and base 9 for fall curl-ups. For each of these bases, the focal (remainder 0) residual is also negative for all four events meaning that there is no evidence of intentional counting by these bases. Note further that each of these bases is odd. In this instance, the just below outcome is even half of the time. By contrast, for even bases, the just below outcome is always odd.

An interesting question is whether those who end at just below outcomes do so as a result of striving for the outcome one repetition higher but simply are not able to achieve that outcome. If that achievement gap occurs because the student has reached his or her performance boundary, then that student's performance may be higher than those who did not attempt to obtain the focal outcome. Difference between means tests confirm that just below performers dominate middle performers rather than the reverse, especially for push-ups.

Erfle and Gelbaugh (2013) and Erfle (2014b) established that middle school children count by 5 and 10 more often than would be expected based on random processes and that those who did count by 5 and 10 outperformed their non-focal peers. This chapter has provided modest evidence that this extends to other bases as well. This chapter also provides a simple way to calculate how intentional such counting structures are. It has also provided evidence that just below skip counting performers may well have been skip counting students who simply reached their performance boundary.

Figure 1. Performance of 10206 Middle School Students on Two Events at Two Assessment Times

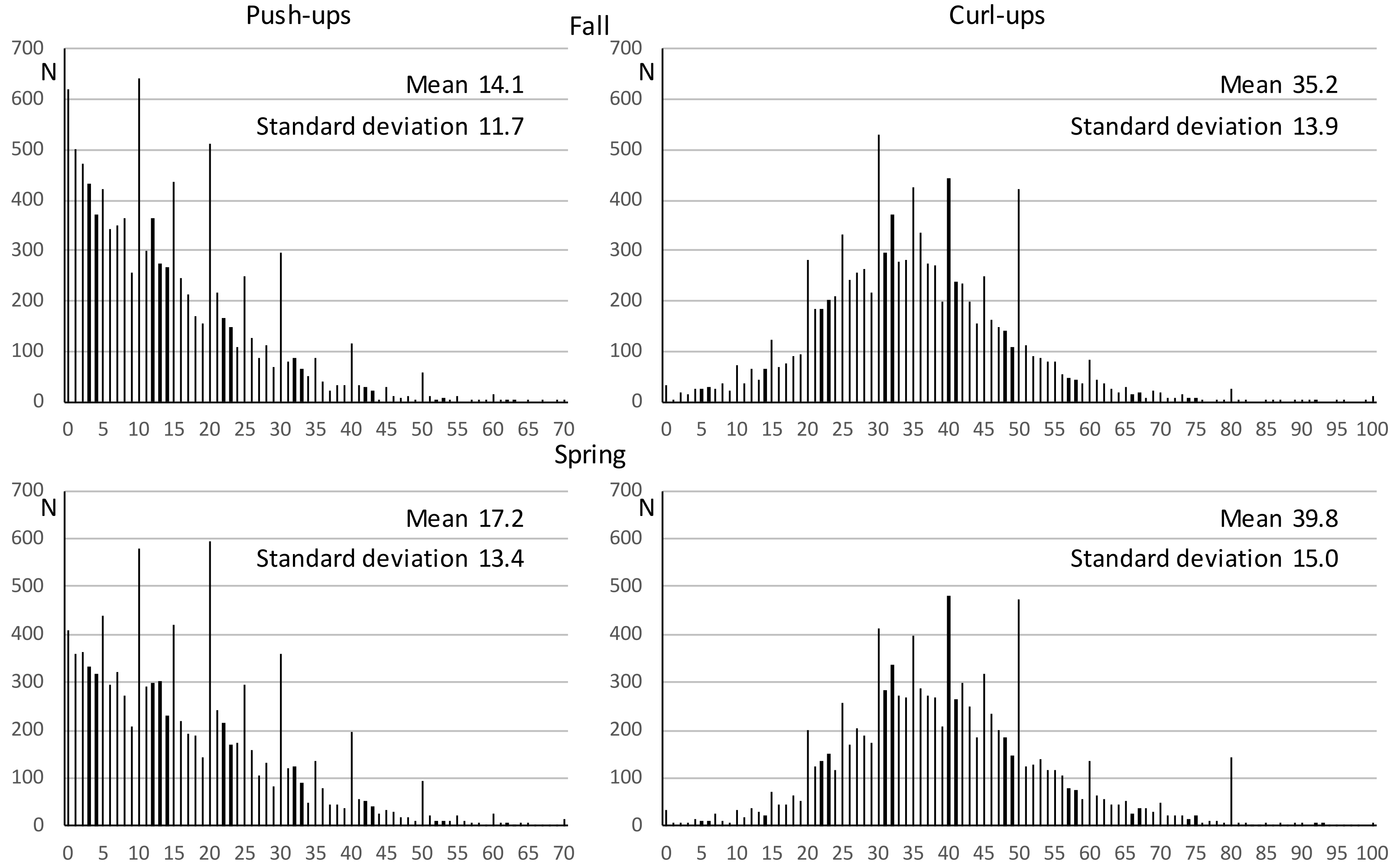
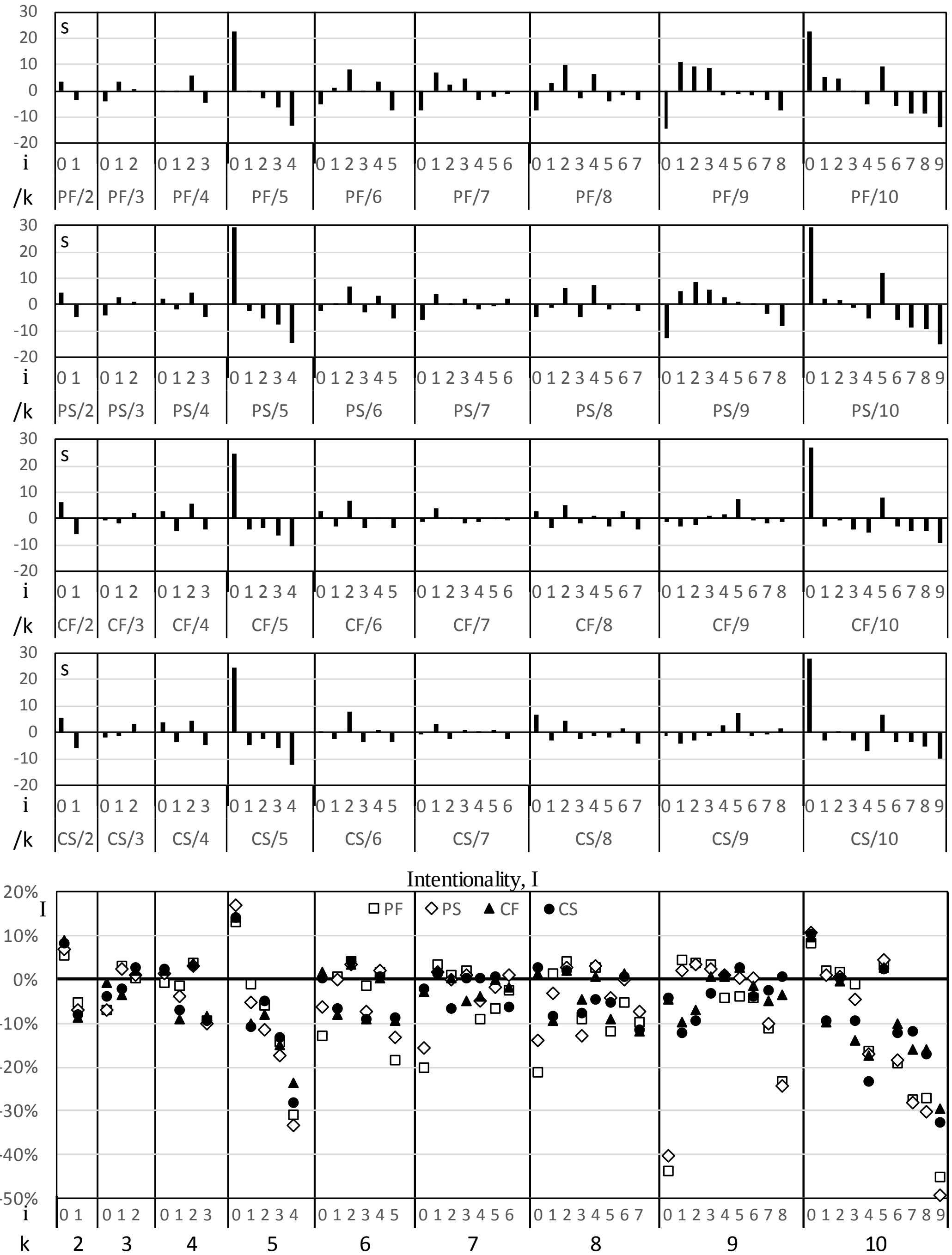


Figure 2. Proclivity Analysis of Counting by Various Bases, k, For Push-ups (P) and Curl-ups (C) in Fall (F) and Spring (S)

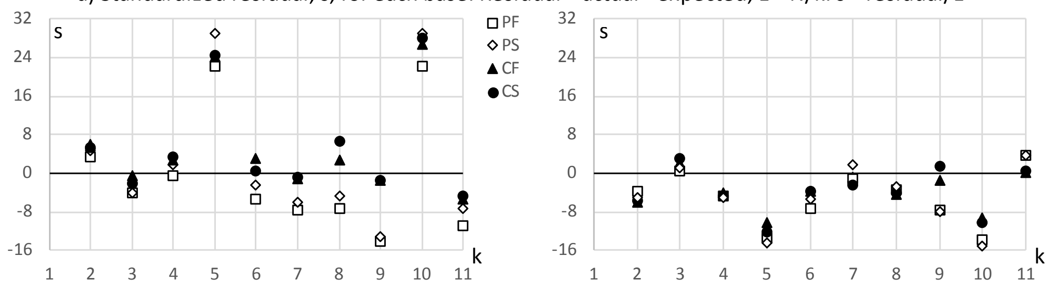


Residual, r = Actual - Expected, E, E = N/k. Intentionality, I, estimates the percentage of individuals intentionally obtaining (for $r > 0$, $I = r/[N - E]$) or avoiding (for $r < 0$, $I = r/E$) a given remainder group. Standardized residual for remainder group i, s_i, $s_i = r_i/E^{0.5}$ assuming uniform remainders upon division by counting base k. All Chi-Square except CF3 are significant at $p < .001$, $\chi^2_{CF3} = 8.9$, $p = .012$. N = 9,345.

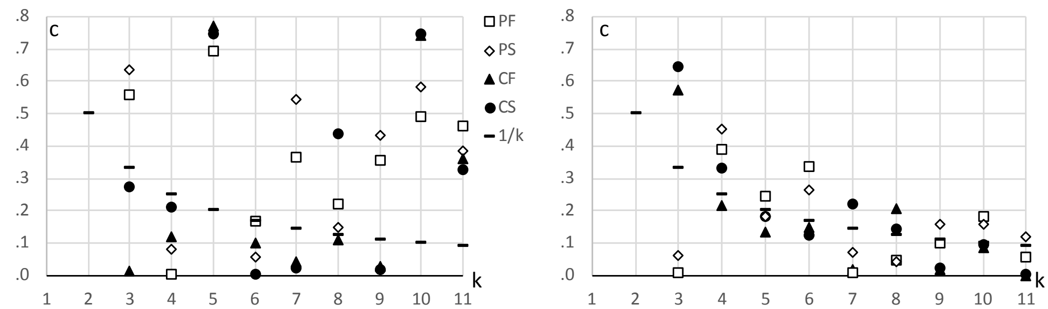
Figure 3. Proclivity of Focal and Just Below Counting by Base k from k = 2 to k = 11 on Four Events

Focal Counting (remainder = 0) | Just Below Focal Counting (remainder = k-1)

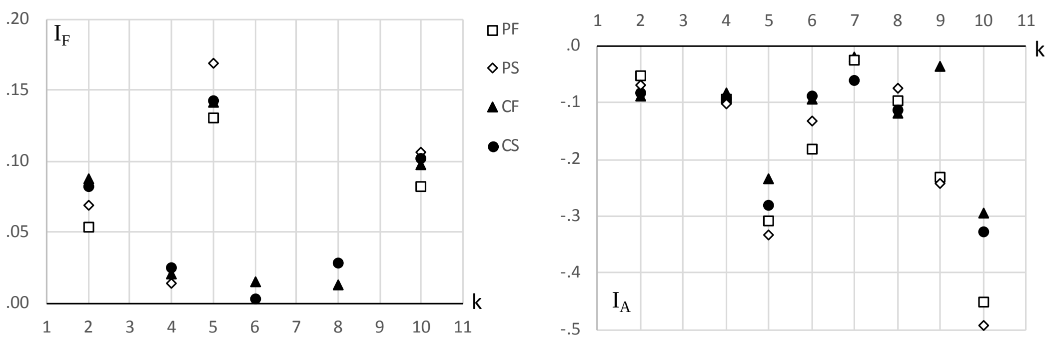
a) Standardized residual, s, for each base. Residual = actual - expected, $E = N/k$. $s = \text{residual}/E^{0.5}$



b) Contribution to Chi-Square from this remainder for each base

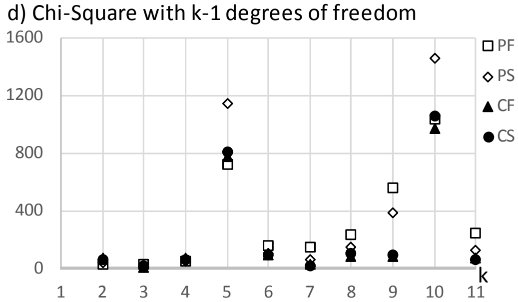


c) Intentionality to obtain Focal, I_F, or avoid Just Below, I_A, for each base



$I_F = (N_{Focal} - E)/(N - E)$ | $I_A = (N_{Just\ Below} - E)/E$

d) Chi-Square with k-1 degrees of freedom



Standardized residuals:

Above 0 -- greater than expected proclivity

Below 0 - less than expected proclivity

Chi-square contribution:

Above 1/k is an important deviation

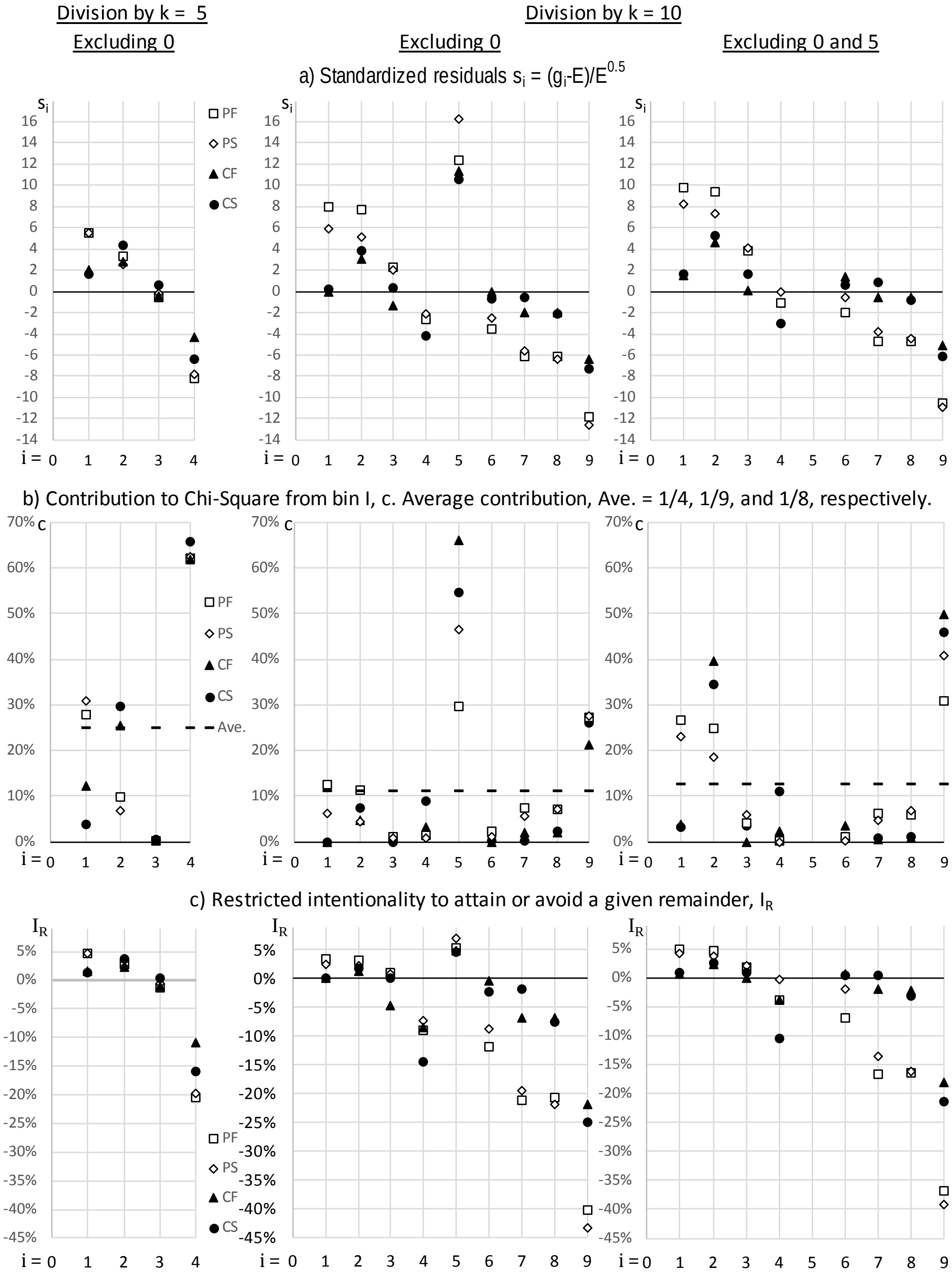
Below 1/k is an unimportant devation

Intentionality is highest for 5 and 10:

Intentionality to avoid just below is higher than intentionality to attain focal

Note. All Chi-Square except CF3 are significant at $p < .001$, Chi-Square$_{CF3} = 8.9$, $p = .012$. N = 9,345.

Figure 4. Proclivity of Last Digits Other than 0 or 5 on Four Events



Note. F = fall, S = spring, C = curl-ups, P = push-ups. Let g_0 , ..., g_{k-1} be the actual number in remainder groups 0 to k-1, respectively. Then $N = g_0 + ,..., + g_{k-1}$. The total nonzero entries. $N_{nz} = N - g_o$ when 0 is excluded upon division by k = 5 or k = 10 and $N_{nz,5} = N - g_0 - g_5$ when 0 and 5 are excluded upon division by 10. The expected number in a nonzero bin, E, is $E = N_{nz}/4$ for k = 5 and $E = N_{nz}/9$ for k = 10 and $E = (N_{nz,5})/8$ when 0 and 5 are excluded upon division by 10. The nonzero residual for bin i is $r_{nz,i} = g_i - E$. If $r_{nz,i} > 0$, the intentionality to achieve remainder i equals $I_R = r_{nz,i}/(N_{nz} - E)$ or $I_R = r_{nz,i}/(N_{nz,5} - E)$ depending on whether 0 or 0 and 5 are excluded. If $r_R < 0$ then the intentionality to avoid a nonzero remainder outcome is $I_R = r_{nz}/E$. All chi-square tests are significant at $p < .001$.

Figure 5. Mean Performance and 95% CI on Six Events by Remainder Group *i* for Counting by Various Bases k for Curl-ups (C) and Push-ups (P) in Fall (F) and Spring (S)

a) Curl-ups

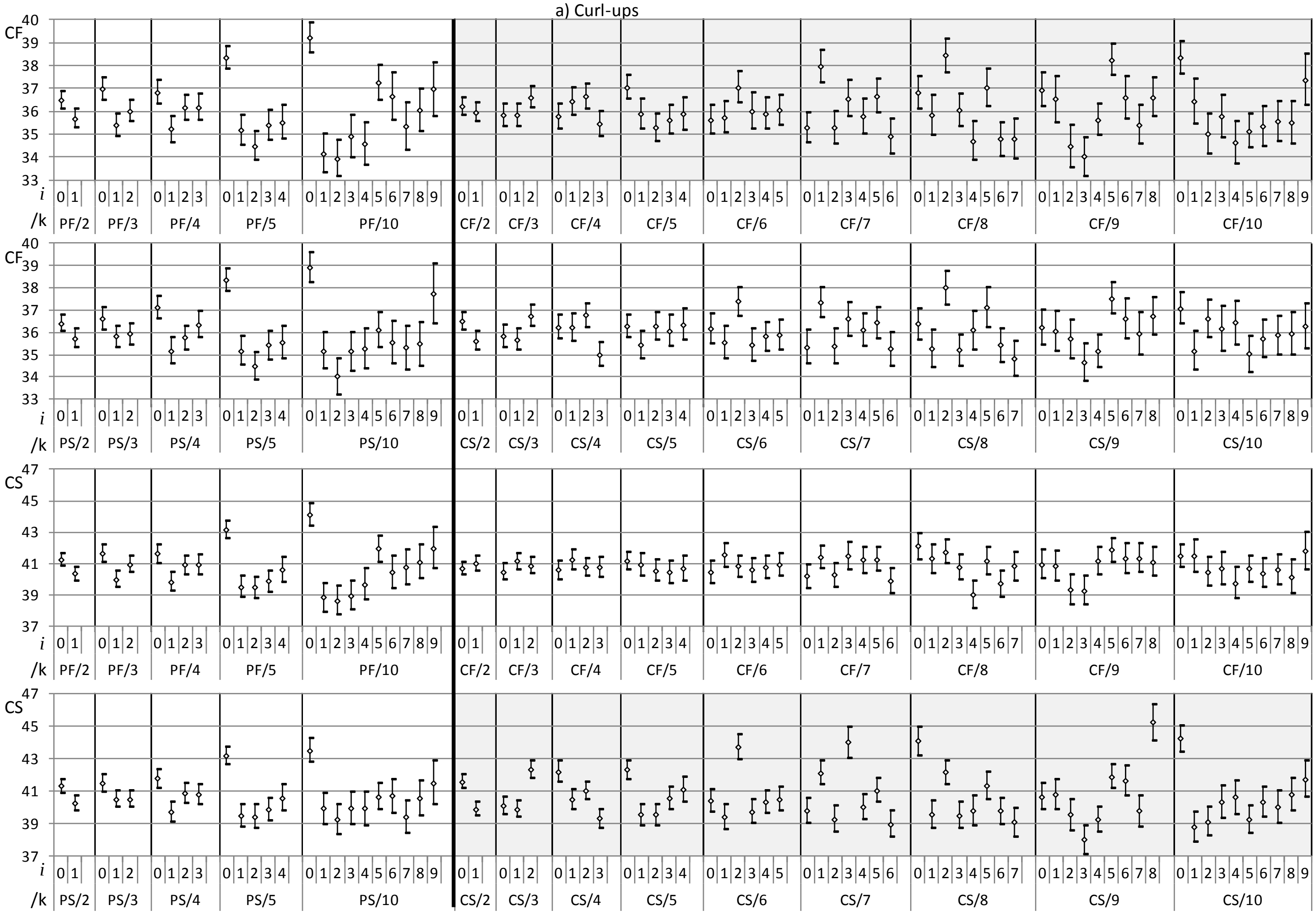


Figure 5 (continued)

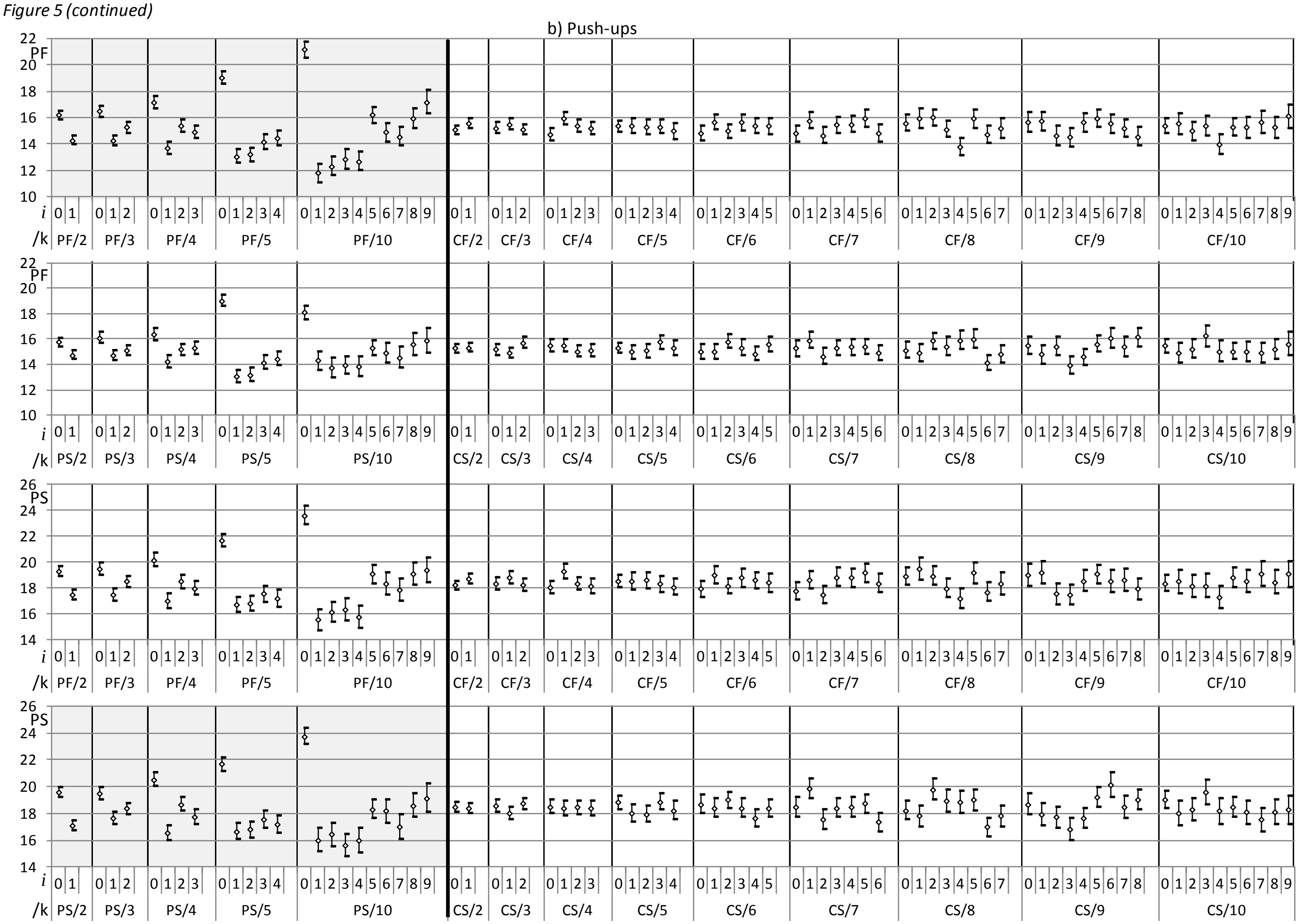


Figure 5 (continued)

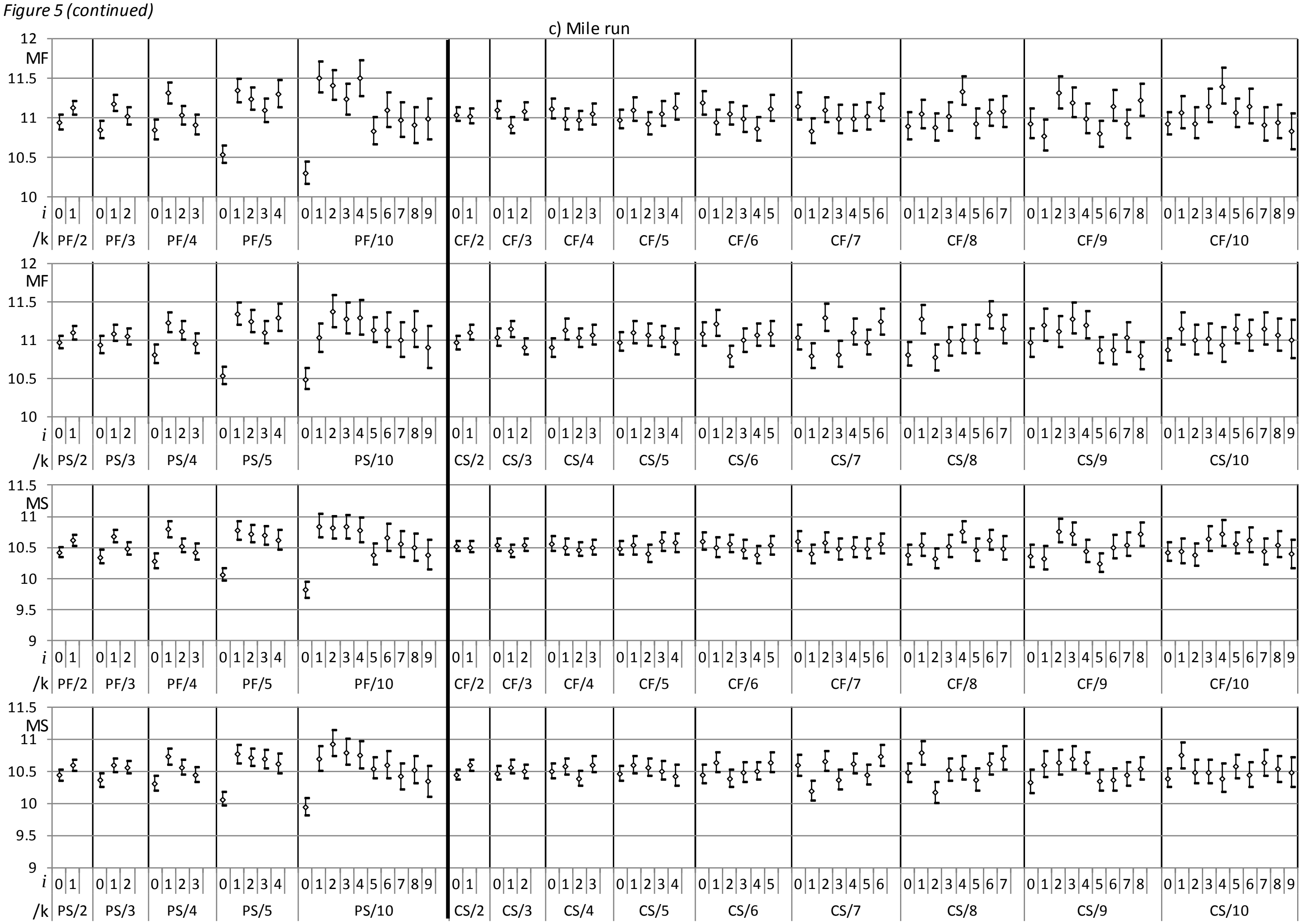


Figure 6. Relative Physical Fitness Performances Using a Three-Remainder Partition for Various Bases, k

a) $Mean_{Focal}$ - $Mean_{Just\ Below}$ with 95% CI on mean difference for three activities at two times

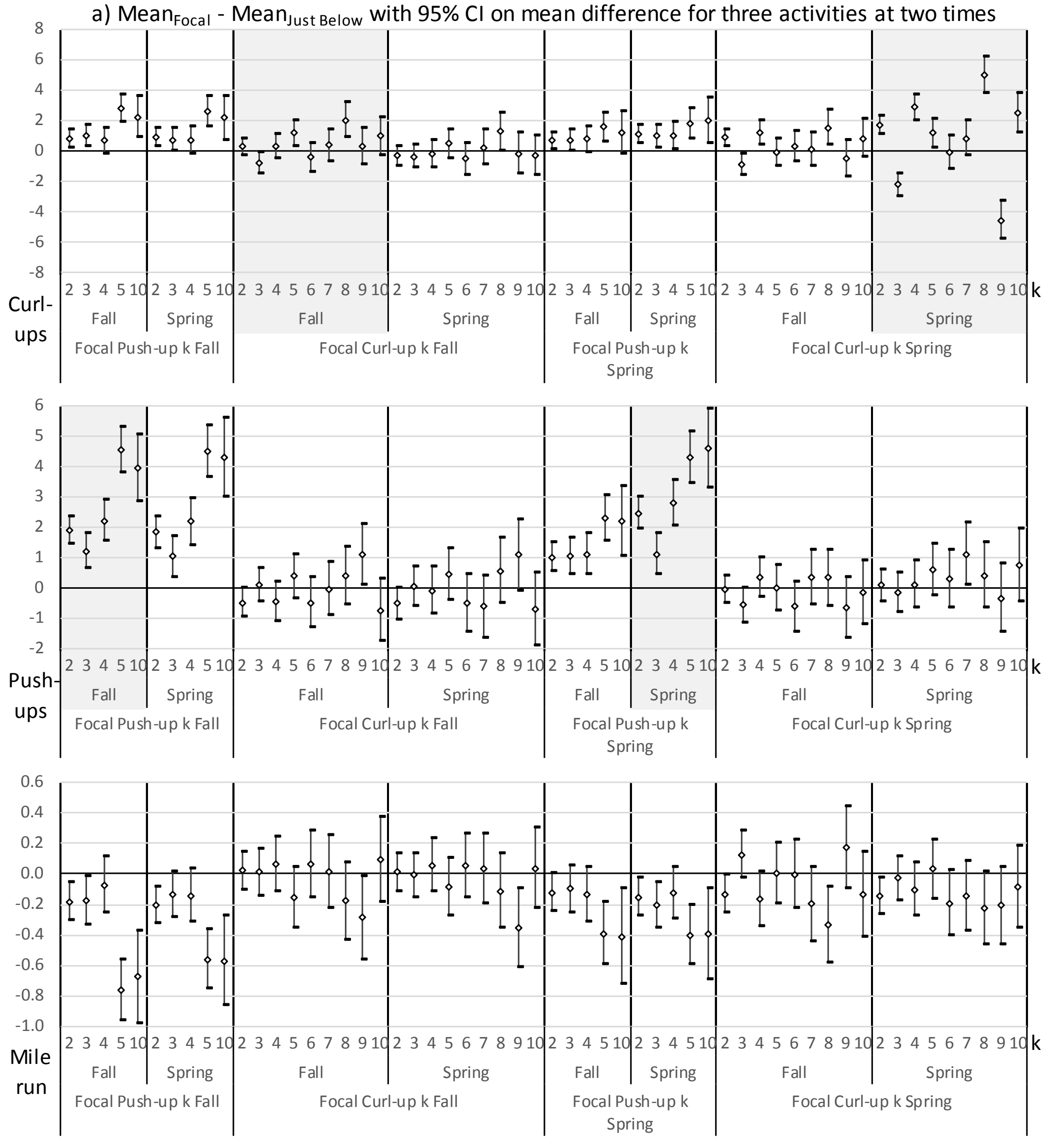


Figure 6 (continued)

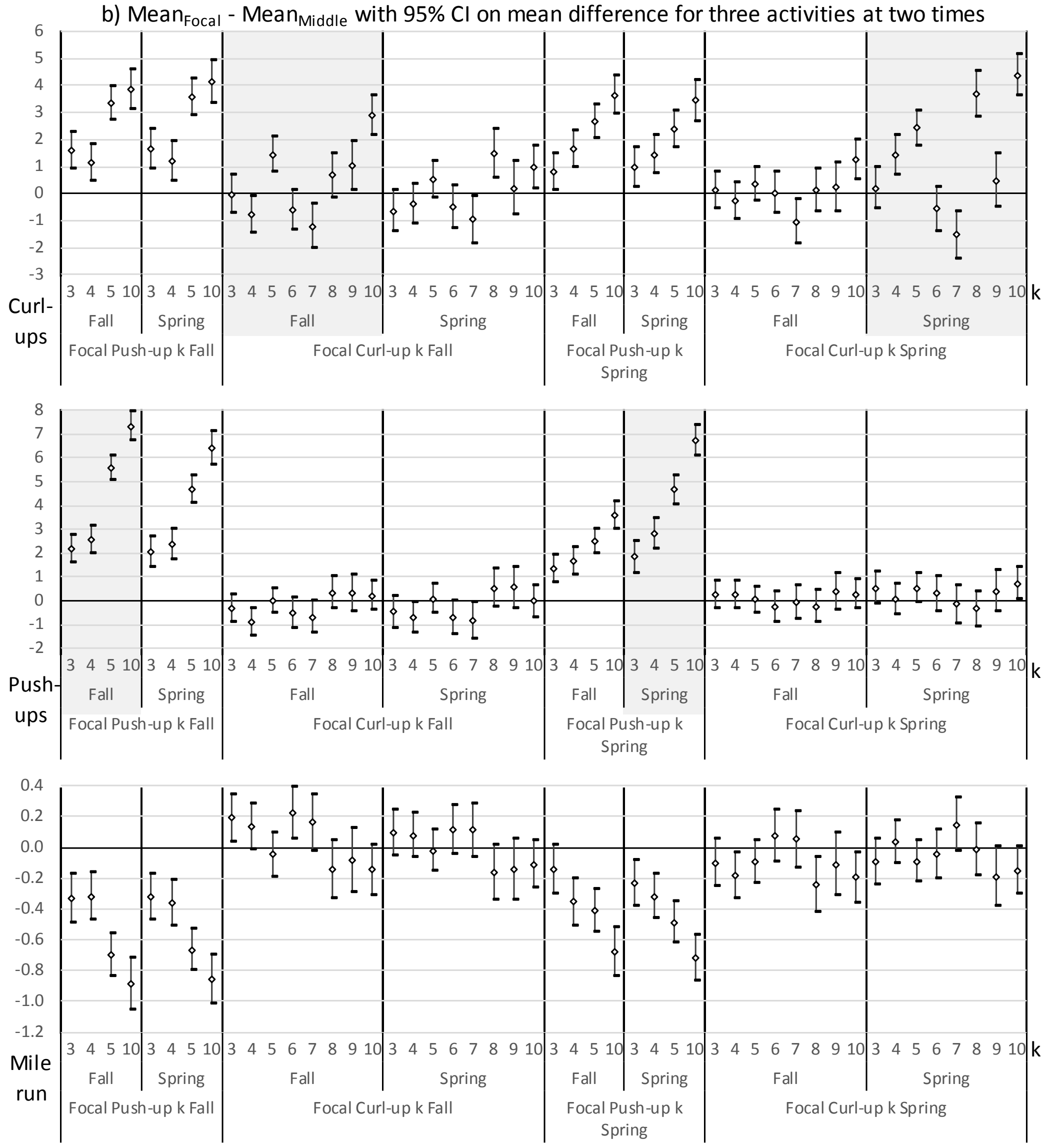
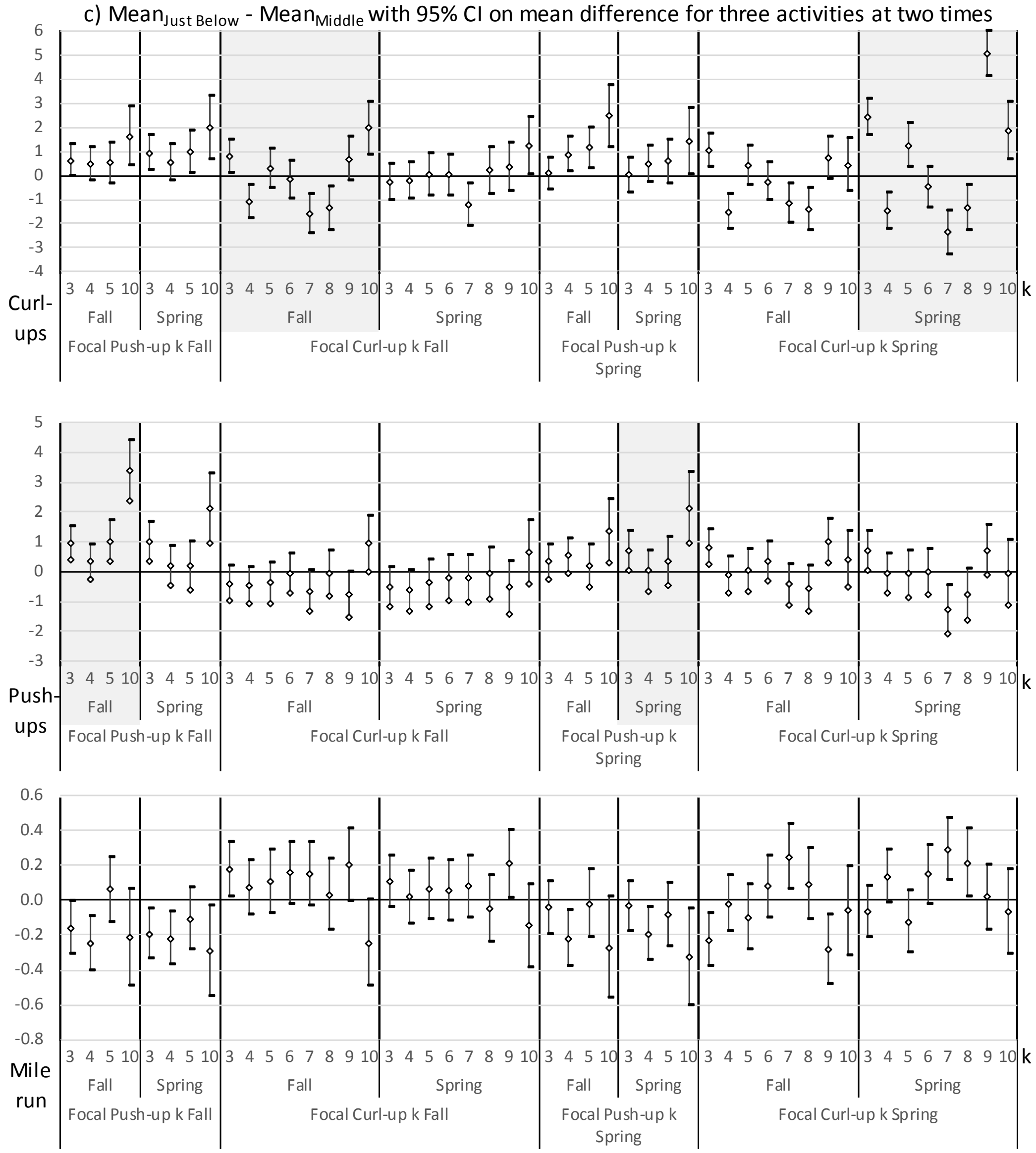


Figure 6 (continued)



Note. Own-event focal performance differences are shaded in each activity panel, all other portions of each panel examine cross-event mean differences. Each three-remainder partition includes a Focal bin comprised of students having remainder 0 when the event performance is divided by k, the Just Below bin has remainder k-1 when the event performance is divided by k and the Middle bin (for k > 2) has remainders 1 to k-2 when the event performance is divided by k. Sample restricted to 9,345 students performing at least one curl-up and one push-up at both assessments.

Table 1. *Testing Even/Odd and Bottom Half/Top Half Independence of Counting by Base 10*

Event:	Fall curl-ups, CF			Spring curl-ups, CS			Fall push-ups, PF			Spring push-ups, PS		
Even/odd	g = CF/2			g = CS/2			g = PF/2			g = PS/2		
g	0	1	Total	0	1	Total	0	1	Total	0	1	Total
Remainder group g = Event/5: 0 N	1,755	1,171	2,926	1,792	1,143	2,935	1,623	1,218	2,841	1,825	1,306	3,131
s 0-4	**4.1**	**-4.5**	**24.4**	**5.1**	**-5.6**	**24.7**	**3.3**	**-3.5**	**22.5**	**3.7**	**-4.0**	**29.2**
1 N	840	843	1,683	819	846	1,665	756	1,092	1,848	762	1,006	1,768
s 0-4	**-2.5**	**2.7**	**-4.3**	**-2.7**	**3.0**	**-4.7**	**-7.0**	**7.3**	**-0.5**	**-6.0**	**6.4**	**-2.3**
s 1-4	**-1.1**	**1.1**	**2.0**	**-1.0**	**1.0**	**1.6**	**-5.9**	**6.0**	**5.5**	**-4.7**	**4.8**	**5.4**
2 N	931	786	1,717	951	824	1,775	1,082	677	1,759	982	672	1,654
s 0-4	**-0.1**	**0.1**	**-3.5**	**-0.3**	**0.3**	**-2.2**	**5.1**	**-5.4**	**-2.5**	**3.3**	**-3.5**	**-5.0**
s 1-4	**1.4**	**-1.4**	**2.8**	**1.6**	**-1.6**	**4.3**	**6.4**	**-6.4**	**3.3**	**4.7**	**-4.8**	**2.5**
3 N	785	804	1,589	776	848	1,624	680	923	1,603	652	892	1,544
s 0-4	**-2.7**	**2.9**	**-6.5**	**-3.5**	**3.8**	**-5.7**	**-5.7**	**6.0**	**-6.2**	**-6.0**	**6.5**	**-7.5**
s 1-4	**-1.4**	**1.4**	**-0.4**	**-1.8**	**1.8**	**0.5**	**-4.7**	**4.7**	**-0.6**	**-4.8**	**4.9**	**-0.2**
4 N	772	658	1,430	717	629	1,346	781	513	1,294	775	473	1,248
s 0-4	**-0.2**	**0.2**	**-10.2**	**-0.4**	**0.4**	**-12.1**	**3.8**	**-4.0**	**-13.3**	**4.2**	**-4.5**	**-14.4**
s 1-4	**1.1**	**-1.2**	**-4.4**	**1.2**	**-1.2**	**-6.4**	**4.9**	**-4.9**	**-8.2**	**5.5**	**-5.6**	**-7.8**

Note: For upper portion of Table 1, shaded values are in the bottom half of division by 10 (g = 0 - 4).

0-4 Total	5,083	4,262	9,345	5,055	4,290	9,345	4,922	4,423	9,345	4,996	4,349	9,345
s 1 × 2	**6.0**	**-6.0**		**5.6**	**-5.6**		**3.7**	**-3.7**		**4.7**	**-4.7**	
Chi-square	2 × 5	1 × 2	1 × 5	2 × 5	1 × 2	1 × 5	2 × 5	1 × 2	1 × 5	2 × 5	1 × 2	1 × 5
Value	**66.5**	**72.1**	**773.7**	**100.0**	**62.6**	**813.5**	**278.6**	**26.6**	**727.0**	**244.6**	**44.8**	**1145**
p	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**
1-4 Total	3,328	3,091	6,419	3,263	3,147	6,410	3,299	3,205	6,504	3,171	3,043	6,214
s 1 × 2	**2.1**	**-2.1**		1.0	-1.0		0.8	-0.8		1.1	-1.1	
Chi-square	2 × 4	1 × 2	1 × 4	2 × 4	1 × 2	1 × 4	2 × 4	1 × 2	1 × 4	2 × 4	1 × 2	1 × 4
Value	**12.8**	**8.8**	**30.9**	**16.4**	2.1	**62.4**	**245.4**	1.4	**109.3**	**199.6**	2.6	**96.3**
p	**.005**	**.003**	**<.001**	**<.001**	.147	**<.001**	**<.001**	.244	**<.001**	**<.001**	.104	**<.001**

Bottom/Top	g = CF/10		g = CS/10		g = PF/10		g = PS/10	
g	0 - 4	5 - 9	0 - 4	5 - 9	0 - 4	5 - 9	0 - 4	5 - 9
Remainder group g = Event/5: 0 N	1,755	1,171	1,792	1,143	1,623	1,218	1,825	1,306
s 0-4	**3.9**	**-4.3**	**4.3**	**-4.8**	**-1.2**	**1.4**	-0.3	0.3
1 N	843	840	846	819	1,092	756	1,006	762
s 0-4	**-2.5**	**2.8**	**-2.4**	**2.6**	**0.1**	**-0.2**	-1.0	1.1
s 1-4	**-1.2**	**1.2**	-0.9	1.0	-0.3	0.4	**-1.1**	**1.3**
2 N	931	786	951	824	1,082	677	982	672
s 0-4	**-0.2**	**0.2**	**-0.9**	**1.0**	**1.4**	**-1.7**	0.4	-0.5
s 1-4	**1.2**	**-1.2**	0.7	-0.7	1.0	-1.2	**0.3**	**-0.4**
3 N	804	785	848	776	923	680	892	652
s 0-4	**-2.2**	**2.4**	**-1.6**	**1.8**	**-0.7**	**0.8**	-0.4	0.5
s 1-4	**-0.9**	**0.9**	-0.1	0.1	-1.1	1.3	**-0.5**	**0.6**
4 N	772	658	717	629	781	513	775	473
s 0-4	**-0.3**	**0.4**	**-0.9**	**1.0**	**0.7**	**-0.8**	1.6	-1.9
s 1-4	**0.9**	**-1.0**	0.4	-0.4	0.3	-0.4	**1.5**	**-1.8**

Note: For lower portion of Table 1, shaded values are even last digits.

0-4 Total	5,105	4,240	5,154	4,191	5,501	3,844	5,480	3,865
s 1 × 2	**6.3**	**-6.3**	**7.0**	**-7.0**	**12.1**	**-12.1**	**11.8**	**-11.8**
Chi-square	2 × 5	1 × 2	2 × 5	1 × 2	2 × 5	1 × 2	2 × 5	1 × 2
Value	**58.6**	**80.1**	**63.4**	**99.2**	**11.0**	**293.8**	9.4	**279.1**
p	**<.001**	**<.001**	**<.001**	**<.001**	**.027**	**<.001**	.053	**<.001**
1-4 Total	3,350	3,069	3,362	3,048	3,878	2,626	3,655	2,559
s 1 × 2	**2.5**	**-2.5**	**2.8**	**-2.8**	**11.0**	**-11.0**	**9.8**	**-9.8**
Chi-square	2 × 4	1 × 2	2 × 4	1 × 2	2 × 4	1 × 2	2 × 4	1 × 2
Value	**9.3**	**12.3**	3.1	**15.4**	5.9	**241.0**	**9.1**	**193.3**
p	**.026**	**<.001**	.377	**<.001**	.117	**<.001**	**.028**	**<.001**

Note: Boldfaced chi-square tests and their assiciated standardized residuals, s, are significant at the p < .05 level based on a 2-tailed test.

Table 2. *Percent of Statistically Significant Difference Between Means Tests using the Three-Remainder Partitions Analyzed in Figure 6*

	Figure 6 Panel:		a) Focal versus Just Below				b) Focal versus Middle				c) Just Below versus Middle		
Focal performance activity	Activity	Own or Cross	Focal dominates Just Below	Just Below dominates Focal	No significant difference	*N-Panel a*	Focal dominates Middle	Middle dominates Focal	No significant difference	*N-Panels b&c*	Just Below dominates Middle	Middle dominates Just Below	No significant difference
Push-ups	Curl-ups	Cross	75%	0%	25%	*20*	100%	0%	0%	*16*	50%	0%	50%
	Push-ups	**Own**	**100%**	**0%**	**0%**	***10***	**100%**	**0%**	**0%**	***8***	**62.5%**	**0%**	**37.5%**
		Cross	100%	0%	0%	*10*	100%	0%	0%	*8*	37.5%	0%	62.5%
	Mile run	Cross	65%	0%	35%	*20*	93.8%	0%	6.3%	*16*	50%	0%	50%
	All activities		80%	0%	20%	*60*	97.9%	0%	2.1%	*48*	50%	0%	50%
Curl-ups	Curl-ups	**Own**	**38.9%**	**16.7%**	**44.4%**	***18***	**43.8%**	**18.8%**	**37.5%**	***16***	**37.5%**	**37.5%**	**25%**
		Cross	22.2%	5.6%	72.2%	*18*	18.8%	12.5%	68.8%	*16*	12.5%	25.0%	62.5%
	Push-ups	Cross	5.6%	5.6%	88.9%	*36*	3.1%	9.4%	87.5%	*32*	12.5%	3.1%	84.4%
	Mile run	Cross	13.9%	0%	86.1%	*36*	9.4%	6.3%	84.4%	*32*	6.3%	15.6%	78.1%
	All activities		16.7%	5.6%	77.8%	*108*	14.6%	10.4%	75.0%	*96*	14.6%	16.7%	68.8%
Across focal activities			39.3%	3.6%	57.1%	*168*	42.4%	6.9%	50.7%	*144*	26.4%	11.1%	62.5%

Note. Percent of 56 pairwise comparisons for each activity in Figure 6a (168 total) and 48 pairwise comparisons in Figures 6b and 6c (144 total). Statistically significant comparisons at $p < .05$ level based on a 2-tailed test and seen visually as a 95% confidence interval on the difference between means whisker that does not cross the horizontal axis in Figure 6. Own events in grayscale as in Figure 6.

Table 3. *Proclivity and Mean Performance by Students who Stop at Prime and Not Prime Number Outcomes Using Full Sample and Seven Subsamples*

		Curl-ups								Push-ups							
			Restricted subsamples								Restricted subsamples						
							Trimmed								Trimmed		
		Full sample	Unbroken series	Full not 0	Odd not 5s	Median or above	10%	20%	Largest prime	Full sample	Unbroken series	Full not 0	Odd not 5s	Median or above	10%	20%	Largest prime
									a) Fall								
Sample	LB	0	0	1	1	34	20	24	6	0	0	1	1	12	1	4	6
	UB	201	76	201	201	201	51	46	109	120	55	120	73	120	30	23	73
	Not prime:Prime	156:46	56:21	155:46	37:44	133:35	25:7	18:5	78:26	91:30	40:16	90:30	11:19	84:25	20:10	13:7	50:18
	N in sample	10,206	10,118	9,345	3,091	5,383	8,246	6,457	10,070	10,206	10,164	9,345	3,205	5,129	8,675	6,218	7,381
	N prime	2,135	2,121	1,957	1,924	1,080	1,575	1,222	2,073	3,020	3,012	2,926	2,079	1,041	2,838	1,863	1,691
Proclivity	r_p = $\text{Residual}_{\text{Prime}}$	**-189.1**	**-638.5**	**-181.7**	**244.9**	-41.5	**-228.8**	**-181.7**	**-444.5**	**489.6**	**108.0**	**589.8**	49.2	**-135.4**	-53.7	**-313.3**	**-262.8**
	s_p = $\text{Std.Resid}_{\text{Prime}}$	**-3.9**	**-12.2**	**-3.9**	**6.0**	-1.2	**-5.4**	**-4.8**	**-8.9**	**9.7**	**2.0**	**12.2**	1.1	**-3.9**	-1.0	**-6.7**	**-5.9**
	Intentionality	**-8.1%**	**-23%**	**-8%**	**17%**	-3.7%	**-13%**	**-13%**	**-18%**	**6.4%**	**1.5%**	**8.4%**	4.2%	**-12%**	-1.9%	**-14%**	**-13%**
	Chi square	**19.9**	**203.1**	**20.0**	**78.2**	1.9	**37.2**	**30.1**	**104.6**	**126.0**	**5.6**	**198.5**	3.2	**20.2**	1.5	**69.4**	**48.1**
	Chi square p	**<.001**	**<.001**	**<.001**	**<.001**	.164	**<.001**	**<.001**	**<.001**	**<.001**	**.018**	**<.001**	.072	**<.001**	.222	**<.001**	**<.001**
Performance	Sample mean, M	35.2	34.7	36.1	36.2	45.1	34.7	34.4	35.6	14.1	13.9	15.2	13.5	23.1	12.4	12.3	18.6
	$M_{\text{NotPrime}} - M_{\text{Prime}}$	**1.1**	**0.9**	**1.0**	**1.0**	-0.2	**-0.7**	**-1.8**	0.3	**4.8**	**4.6**	**6.2**	**-1.5**	**1.9**	**5.1**	**1.2**	**4.8**
	ΔM as % of M	**3%**	**3%**	**3%**	**3%**	-0.3%	**-2%**	**-5%**	1%	**34%**	**33%**	**41%**	**-11%**	**8%**	**41%**	**10%**	**26%**
	2-tailed t test p	**.002**	**.006**	**.005**	**.043**	.657	**.002**	**<.001**	.385	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**

Table 3 (continued)

		Curl-ups								Push-ups							
			Restricted subsamples								Restricted subsamples						
							Trimmed								Trimmed		
		Full sample	Unbroken series	Full not 0	Odd not 5s	Median or above	10%	20%	Largest prime	Full sample	Unbroken series	Full not 0	Odd not 5s	Median or above	10%	20%	Largest prime
		b) Spring															
Sample	LB	0	0	1	1	39	23	28	6	0	0	1	1	15	2	5	6
	UB	150	90	150	133	150	58	50	107	132	70	132	101	132	35	28	101
	Not prime:Prime	116:35	67:24	115:35	24:30	89:23	28:8	17:6	77:25	101:32	52:19	100:32	17:24	92:26	23:11	17:7	73:23
	N in sample	10,206	10,157	9,345	3,147	5,141	8,187	6,408	10,118	10,206	10,178	9,345	3,043	5,176	8,467	6,491	7,979
	N prime	2,158	2,149	1,975	1,964	1,079	1,733	1,445	2,132	2,941	2,938	2,736	2,050	886	2,762	1,861	1,804
Proclivity	r_p	**-207.6**	**-529.8**	**-205.5**	**215.7**	23.3	**-86.3**	**-226.7**	**-347.9**	**485.4**	**214.3**	**470.5**	**268.7**	**-254.5**	22.7	-32.2	**-107.6**
	s_p	**-4.3**	**-10.2**	**-4.4**	**5.2**	0.7	**-2.0**	**-5.5**	**-7.0**	**9.8**	**4.1**	**9.9**	**6.4**	**-7.5**	0.4	-0.7	**-2.5**
	Intentionality	**-8.8%**	**-20%**	**-9.4%**	**15%**	0.6%	**-4.7%**	**-14%**	**-14%**	**6.3%**	**2.9%**	**6.6%**	**21%**	**-22%**	0.4%	-1.7%	**-5.6%**
	Chi square	**23.7**	**142.3**	**25.3**	**59.9**	0.6	**5.3**	**41.6**	**64.7**	**126.4**	**23.0**	**129.0**	**97.8**	**72.8**	0.3	0.8	**8.0**
	Chi square p	**<.001**	**<.001**	**0**	**<.001**	.422	**.022**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.598	.379	**.005**
Performance	Sample mean, M	39.8	39.5	40.8	40.1	51.1	39.0	38.8	40.1	17.2	17.0	18.4	16.6	27.4	15.4	14.9	21.2
	$M_{NotPrime}–M_{Prime}$	**1.4**	**1.4**	**1.4**	0.6	**1.9**	**1.4**	**1.1**	**1.1**	**6.4**	**6.2**	**7.4**	0.7	0.7	**7.1**	**4.8**	**3.6**
	ΔM as % of M	**4%**	**3%**	**3%**	1%	**4%**	**4%**	**3%**	**3%**	**37%**	**37%**	**40%**	4%	2%	**46%**	**32%**	**17%**
	2-tailed t test p	**<.001**	**<.001**	**<.001**	.295	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.194	.079	**<.001**	**<.001**	**<.001**

Note. LB = Lower bound. UB = upper bound. M = sample mean. Residual and chi square calculations based on an assumption of random placement across all possible prime and not prime performance outcomes in the subsample range. Intentionality calculations are as defined in the text. Boldfaced cells are significant at $p < .05$ level. Shaded cells exhibit reverse sign from expected proclivity ($residual_{prime} < 0$) or performance ($M_{NotPrime} > M_{Prime}$).

REFERENCES

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology, 39*, 338-366. doi:10.1037/h0080065

Erfle, S. E., & Gamble, A. (2015). Effects of daily physical education on physical fitness and weight status in middle school adolescents. *J Sch Health, 85*(1), 27-35.

Erfle, S. E. (2014a). Analyzing the effects of daily physical education in middle schools on obesity: Evidence from Pennsylvania's active schools program. In R. Todaro (Ed.), *Handbook of physical education research: Role of school programs, children's attitudes and health implications* (pp. 91-108). New York: Nova Science Publishers, Inc.

Erfle, S. E. (2014b). Persistent focal behavior and physical activity performance. *Measurement in Physical Education & Exercise Science, 18*(3), 168-183.

Erfle, S. E., & Gelbaugh, C. M. (2013). Physical activity performance of focal middle school students. *Measurement in Physical Education & Exercise Science, 17*(2), 150-166.

Katz, I., Assor, A., & Kanat-Maymon, Y. (2008). A projective assessment of autonomous motivation in children: Correlational and experimental evidence. *Motivation and Emotion, 32*(2), 109-119. doi:10.1007/s11031-008-9086-0

Mayer, J. D., Faber, M. A., & Xu, X. (2007). Seventy-five years of motivation measures (1930-2005): A descriptive analysis. *Motivation and Emotion, 31*(2), 83-103. doi:10.1007/s11031-007-9060-2

Reviewed by Todd A. Crowder, United States Military Academy