Peers, Pressure, and Performance at the National Spelling Bee

Jonathan Smith*

The College Board – Advocacy and Policy Center

May 2012

Abstract. *This paper investigates how individuals' performances of a cognitive task in a high pressure competition are affected by their peers' performances. To do so, I use novel data from the National Spelling Bee, in which students attempt to spell words correctly in a tournament setting. Across OLS and instrumental variables approaches, I find that when the immediate predecessor is correct, a speller has a 13 to 64 percent greater probability of making a mistake, relative to the predecessor being incorrect. There is no evidence that the effect differs by gender and marginal evidence that it differs by experience.*

Keywords: Competition, Pressure, Peer Effects, Spillovers, Education, Gender Differences

*Email address: jsmith@collegeboard.org.

1. Introduction

There is growing consensus that competitive environments have causal effects on performance, both in the workplace and classroom. Initially, theory and empirics provided guidance on the incentives of tournaments (Lazear and Rosen 1981; Nalebuff and Stiglitz 1983; Bull et al. 1987). A more recent literature has begun to empirically investigate the heterogeneous nature of the response to competition. For example, evidence now suggests substantial gender differences in the way people react to competition (Gneezy et al. 2003; Gneezy and Rustichini 2004; Niederle and Vesterlund 2007; Paserman 2011). But performance in competitive environments may be affected not only by the characteristics of a participant, but also by the participant's peers and their performances in the competition. For example, Brown (2011) shows that peers matter, in that Tiger Woods' entrance into a golf tournament causes other golfers to perform worse, and that peer performance matters, in that the effect is stronger when Woods is performing well.

This paper investigates how individuals' performances are affected by their peers' performances using data from a real world, high stakes competition: The National Spelling Bee. The Bee is a yearly competition in which top grade school students stand on stage, listen to a pronounced word, and orally attempt to provide the correct spelling. Moving one student at a time, those spelling their person-specific word correctly advance in the tournament until a single speller remains and is declared champion.

The novel setting provides a natural experiment to test how performance on a cognitive task can be affected by competitors' performances. Not only does the tournament setup imply that relative performance matters, but the onstage and sequential nature of The Bee allows for

observation of others' performances. The sequencing also allows for the possibility that peers matter differentially, depending on proximity in the ordering.

I begin by testing whether students are more likely to make spelling errors when their immediate predecessors spell correctly, relative to their predecessors spelling incorrectly. To identify the causal impact, I exploit the fact that spelling order is effectively random within any given round, conditional on state residence, and that word difficulty also varies randomly within a round. In ordinary least squares models, I find that when an immediate predecessor spells correctly, a speller's probability of making an error is 13 percent greater than when an immediate predecessor is incorrect. To eliminate the role of common temporal shocks, I also estimate instrumental variables models designed to isolate the role of predecessor ability. As expected, these models produce larger estimates, implying a 64 percent greater probability of making an error.

I then investigate whether spellers other than the immediate predecessor affect a speller's performance. I find evidence that spellers' performances are marginally affected by the result of two spellers before them but not affected by the result of three, four, or five spellers before them. These results imply that more salient peers have a larger effect on a speller's performance. This is consistent with findings in Mas and Moretti (2009) who find that spillovers are more pronounced for workers in close proximity.

Finally, while I find that effects decay across rounds of play and with speller experience, I find no statistically significant differences in effects by gender. Most of the aforementioned literature on gender differences in competitive environments finds females performing relatively worse in such events. However, The Bee substantially differs from previously researched settings. Researchers typically study a physical activity (e.g. Paserman 2011) or a laboratory

experiment (e.g. Gneezy et al. 2003), but I use spelling, a cognitive skill, in The Bee tournament, a selective non-laboratory setting.

This paper's findings have potential implications for group environments in the workplace. The most general interpretation is that own performance on a cognitive task can be negatively affected by peer performance. Most existing research on workplace spillovers understandably relies on physical activities, such as golf (Brown 2011, Guryan et al. 2009), or low-skilled workers, such as cashiers (Mas and Moretti 2009), fruit pickers (Bandiera et al. 2005), or glass workers (Lazear 2000), because they involve somewhat easier to measure tangible tasks. With the notable exception of Brown (2011), these studies do not tend to find negative spillovers. For example, Guryan et al. (2009) find no evidence that randomly assigned golf partners' abilities affect performance, while Mas and Moretti (2009) find that low productivity cashiers positively benefit from the presence of highly productive cashiers. In contrast, my results suggest that, for those who use cognitive skills, being near high performing workers may be detrimental.

This research also implies that in a narrower setting, such as auditions, performances, and competitions, order matters. Many examples of these tournaments are non-cognitive, such as Ms. America pageants, Olympic figure skating, and American Idol, but also include cognitive skills, such as debates and game shows. But perhaps more common work settings that are arguably sequential, such as meetings and presentations, are also relevant.

Finally, the finding that only close peers matter to a speller's performance has potential implications for classroom policy. Perhaps classroom seating arrangements can be a policy maker's tool. Seating arrangements may dictate the order of questioning, grouping, presentations, or oral exams. And while tournaments are not the norm in classrooms, any competitive or high pressure task, such as sequential oral presentations, shares characteristics

with The Bee. Consequently, any spillovers can potentially be mitigated with alternative testing and assignment structures. Moreover, a general lesson is that peer effects in education are likely more complex than recognized in the literature, which may partially explain the mixed findings (e.g., Hoxby 2000, Angrist and Lang 2004, Lavy et al. 2009). Peer effects in schools operate through extensive repeated interactions, not present in the Bee, but may reflect peer learning and peer competition, the latter of which is the focus of the Bee.

2. The Spelling Bee: History, Rules, and Regulations

2.1 History

The modern day National Spelling Bee dates back to 1925 and has been held every year since with the exception of 1943-1945 due to World War II.[1] The first event was held in 1925 by the Louisville Courier-Journal in Kentucky and taken over by E.W. Scripps Company in 1941. The first event consisted of nine contestants and the winning word was "gladiolus." The latest competition in this paper, in May of 2008, consisted of 288 contestants and the winning word was "guerdon."

The popularity of the contest, now held in Washington, D.C., is one reason for the growth in contestants and their respective regional sponsors. Over 10 million students compete at the local level. In turn, the popularity caused ESPN, the nation's leading cable sports channel, to air parts of the competition since 1994. In 2006, while the earlier portion of the competition remained on ESPN, the championship rounds aired on ABC, a major national network, during prime time hours.

2.2 Rules & Format

[1] Ben Bernanke competed in 1965, but did not win.

The exact rules of The Bee have changed over the years but the basic structure remains the same.[2] The main idea of The Bee is to orally spell words correctly until a single champion remains. Spellers all sit on stage and view each others' attempts, one speller at a time. The ordering of spellers is determined first alphabetically by state and then within a state, alphabetically by city (of sponsor). Words given to spellers are unknown in advance other than that they come from *Webster's Third New International Dictionary*.

2.2.1 Eligibility

A contestant partakes in the The Bee by winning a local competition and meeting certain eligibility requirements. There are quite a few eligibility requirements but the basic ones state that a student must be under the age of 15, not past the 8th grade, and must attend a school that is officially registered with the The Bee.[3] Most schools are within the 50 states, but there are several from outside the states, such as Canadian provinces, U.S. territories, and other English speaking countries.

2.2.2 Basic Rules

The Bee begins with the preliminary rounds consisting of an off-stage written round and an on-stage oral round. In the written round a contestant spells 25 words. A speller receives one point for each correctly spelled word. In the oral round, one speller at a time stands on stage and is given a unique word by the pronouncer, which must then be spelled aloud. A correct spelling is worth three points and zero otherwise. The scores from the oral and written rounds are

---

[2] The full set of eligibility requirements and rules for the most recent competition can be found on The Bee's website, http://www.spellingbee.com/contest-rules.

[3] Registered schools does not exclude home schools, which made up 12.5 percent of contestants in 2008.

summed for a maximum of 28 points. The contestants' scores are sorted in descending order and the 100 spellers with the highest scores move past the preliminary rounds.[4]

After the preliminary rounds are the quarterfinals, semifinals, and finals, all of which are on-stage and similar to the aforementioned oral round. However, an incorrect spelling eliminates the contestant from The Bee. Once every speller has attempted to spell one word, a "round" is complete and those who spelled their respective words correctly advance to the next round. When only two or three contestants remain, the pronouncer moves to the "25-word championship section" of the word list. If a single champion does not arise after the 25 words, the remaining contestants are co-champions.

2.2.3 Prizes

Prizes are awarded based on the round of elimination. All contestants that were eliminated in a given round, regardless of order, finish in the same place. The prizes described below are for 2008, but similar payouts can be found in previous years.

All spellers receive a dictionary, wrist watch, and $100 savings bond. The champion of The Bee gets $30,000 in cash, a $5,000 scholarship, $2,500 in U.S. savings bonds, references and a test preparation kit from *Encyclopedia Britannica* valued at $3,800. Then cash awards for second to seventh place are $12,500, $7,500, $3,000, $2,500, $2,000, $1,500, respectively. Those eliminated in round nine to eighth place receives $1,000 and those eliminated in rounds eight, seven, six, five, and four earn $750, $500, $400, $250, $175, $125, respectively. Those eliminated in the preliminary rounds get $50.

[4] If there are ties that would result in more than 100 spellers, those spellers are eliminated. Also, the threshold was initially set at 90 spellers.

3. Data

I compiled data from the 1996-2008 competitions. The data include basic identifying information, such as the speller's name and the city and state of the sponsor. Using the name and state, I match across years to impute experience, which is the number of years the speller has appeared in The Bee.[5] Though the written test scores are not available, the data include complete information on performance in the oral rounds. For each round, both the correctly spelled word given by the pronouncer and the spelling given by the contestant are recorded. Speller order is also observed, allowing me to create indicators for whether immediate and more distant predecessors are correct, and to measure a speller's relative position within that round.

Speller demographic information, beyond geography, is only available for the 2004-2008 competitions. In these years, there exist short biographies of the spellers, complete with photographs, which aid in the construction of several variables including age, grade, and whether homeschooled. A student is determined to be homeschooled if the school name in the biography contains the word or phrase "home." Gender is determined from the text, which usually contains the pronouns "he" or "she", and if not, is evident from the photo and name. Race/ethnicity is assigned using the photo and name of the speller and put into one of the following categories: White, Black, Hispanic, South Asian and Other. Students categorized as Other include those from the Middle East or Pacific Islands, as well as those of mixed race/ethnicity.[6]

---

[5] Experience is set to missing in 1996. Almost all spellers have one or two years of experience, so very few spellers in the next few years may have an inaccurate measure of experience.

[6] This assignment process involves some subjectivity, so a second person also performed the task (for 2005-2008 since the 2004 data became unavailable online). The assignments match one

The analysis primarily uses data from the on-stage rounds of the 2004-2008 competitions because of the availability of student demographic variables. Own and peer characteristics are central to implementing and validating the empirical models. In addition, the structure and rules of The Bee were somewhat different prior to 2004.[7] The preliminaries (round 1), quarterfinals (round 2), and semifinals (round 3 to finals) are included in the analysis. The final rounds, when two or three spellers remain, are excluded because there are too few observations for much of the coming analysis.

Table 1 shows statistics on error rates overall and by round. On average, spellers misspell the word 29.3 percent of the time. The error rate falls from 27.3 percent in the preliminary rounds to 20.8 percent in the quarterfinals, perhaps due to the elimination of less able spellers. Increasing word difficulty may explain why the error rate then rises to 35.7 percent across the semifinal rounds.

Table 2 displays summary statistics for spellers (1,386) and speller-rounds (2,906) included in the analysis. The typical speller is about 13 years old, in seventh grade, and equally likely to be male as female. The mean experience level is 1.27 years, with a maximum of 5 years. A disproportionate number of spellers (12 percent) are homeschooled.[8] South Asians are also

---

another for 92.2 percent of the 1,121 spellers, and the findings are not sensitive to which assignment is used.

[7] A description of how The Bee has changed over the years is in the web appendix, along with summary statistics and results using 1997-2003 data. In general, evidence for negative spillovers from peer performance is weaker in the earlier period.

[8] Though inferred from the biographies, the aggregate rate matches statistics reported by The Bee.

overrepresented, while blacks are highly underrepresented. Finally, nearly half of the observations at the speller-round level are from the preliminaries.

4. Estimation and Identification

To investigate whether spellers affect other spellers' performances, I test whether the probability of spelling correctly is affected by predecessors and other factors using the following general specification:

$$\Pr(y_{i,rt} = 1) = \alpha \cdot y_{i-1,rt} + X_i\beta + Z_{i,t}\gamma + S + T + \varepsilon_{i,rt}$$

with the following definitions:

- $y_{i,rt}$: Equals 1 if speller *i* in round *r* and year *t* spells the word correctly, 0 otherwise.
- $y_{i-1,rt}$: Equals 1 if speller *i*-1 in round *r* and year *t* spells the word correctly, 0 otherwise.
- X_i : Speller *i* characteristics that do not change over years (e.g. gender and race/ethnicity)
- $Z_{i,t}$: Speller *i* characteristics that do change over years (e.g. age, grade, whether homeschooled, and experience).
- S : State fixed effects for the speller's state.
- T : Time fixed effects (at round-year level).
- $\varepsilon_{i,rt}$: An i.i.d. error term.

I start by estimating the above equation with a linear probability model.[9] That is, $\Pr(y_{i,rt} = 1) = y_{i,rt}$.[10] The estimate of α captures systematic differences in spelling performance when spellers follow a predecessor who spells correctly, relative to one who spells incorrectly. For this to be interpreted as causal, the unobserved determinants of spelling correctly must be uncorrelated with the predecessor's performance, conditional on observables.

One possible unobserved determinant is word difficulty.[11] Prior to the competition, judges group words by round, so that average word difficulty systematically increases across rounds. In order to account for this, all specifications include round-year fixed effects, so that identification comes from variation within a round and year. Variation in word difficulty at this level is in fact random.

Another possible unobserved determinant is speller ability. Just as word difficulty increases, underlying ability systematically increases across rounds as less able spellers are eliminated. While the round-year fixed effects eliminate this source of correlation in performance across subsequent spellers, speller order within a round is not randomly assigned. Within a round, students are first organized alphabetically by state, and then alphabetically by the sponsor's city. To mitigate the concern that students from the same state are likely to follow one another and to be similar in terms of unobserved ability, I include state fixed effects. Moreover, as rounds

[9] The first speller in all rounds is excluded when estimating the effect of the immediate predecessor, as they do not have a predecessor within that round.

[10] I test the sensitivity of the results with a probit model.

[11] My attempts to create informative measures of word difficulty were not successful. For example, word length does not help to predict whether a speller spells correctly.

progress and spellers are eliminated, the person next to any given speller continually changes, so the ordering with respect to unobserved ability is more likely to approximate random assignment.

A final threat to causal interpretation is common time effects that lead to serial correlation in performance. For example, if fatigue increases with time elapsed in the competition, performance may decline across spellers within a round. Alternatively, competitive pressure may rise or fall. To control for these types of performance trajectories, I control for speller position within a round.

In the presence of negative peer performance effects, each of the confounders discussed above (unobserved word difficulty, ability, and time effects) would tend to bias the estimated impacts toward zero. That is, these should induce a positive correlation in performance across consecutive spellers. Thus, adding the controls intended to eliminate these sources of bias should lead to either unchanged or more negative estimates of the predecessor spelling correctly. I test this by starting with a sparse specification that includes only round-year fixed effects, and then progressively add state fixed effects and position within round. I then also add the full set of student characteristics (gender, experience, race/ethnicity, age, grade and whether homeschooled). The results should be unchanged if consecutive spellers' characteristics are in fact uncorrelated, conditional on the controls.

In addition to the effects of an immediate predecessor's performance, that is, $y_{i-1,rt}$ is the variable of interest, I also test alternatives, such as multiple predecessors' performances. I also test for differential effects by round, gender, and experience. The same identification applies to estimation of the effects of several predecessors. That is, from a speller's perspective, the two, three, or more immediate predecessors are randomly determined and the difficulty of their words is also random, conditional on position in round and round-year and state fixed effects.

4.1 Instrumental Variables

A remaining threat to validity is common shocks to proximate spellers that are local or nonlinear in nature. For example, someone in the audience may cough through several spellers' attempts. Or, spellers may feel more pressure spelling early, and again as the round nears its end. Such shocks are likely to be positively correlated across spellers, suggesting the OLS estimates may be downwardly biased even conditional on the included temporal controls.

I solve this problem using an instrumental variables approach. I treat predecessor performance as endogenous, and use the predecessor's characteristics as instruments. I instrument "speller *i*-1 correct" with speller *i*-1's gender, age, grade, experience, whether homeschooled, and race/ethnicity. These characteristics are potentially predictive of speller *i*-1's performance, which is tested in the first stage.

The instrumental variables approach requires that none of speller *i*-1's characteristics affect speller *i*'s performance directly, but rather indirectly through speller *i*-1's performance. This assumption would not hold if, for example, spellers' confidence or focus varies depending on whether they are following someone of the same gender, regardless of how that person performs. When including the full set of predecessor characteristics as instruments, I test validity using Sargan's overidentification test. I also estimate a just identified model with speller *i*-1's experience as the sole instrument, since this is perhaps the most likely to satisfy the exclusion restriction *a priori*.

In addition to estimating standard IV models, I also estimate the two-step method used by Guryan, Kroft and Notowidigdo (2009) (the GKN method). I first predict speller performance, by regressing whether a speller is correct on her characteristics and fixed effects for round and

state, excluding the year of the speller in question. I then regress whether speller *i* is correct on the predicted performance of speller *i*-1, conditional on the usual control variables.

These approaches also facilitate a placebo test, whereby speller *i*'s performance is regressed on speller *i*-1's predicted performance, and a test of random assignment and the validity of the instruments, whereby speller *i*'s predicted performance is regressed on speller *i*+1's predicted performance.

When comparing the OLS and IV results, it is important to recognize that results may differ for two reasons. First, the IV estimates may be more negative if common shocks induce positive serial correlation. Second, the OLS model estimates the average response to peer success, regardless of whether the peer correctly spelled a relatively difficult or easy word. The IV model isolates success that is orthogonal to word difficulty, so it estimates the response to following a speller who spelled a word of typical difficulty correctly due to competence. Spellers may respond more strongly to this form of peer success than to peer success attributable to the luck of the word draw.

5. Results

5.1 OLS

Table 3 contains OLS results on whether the immediate predecessor being correct affects the speller's performance, relative to the predecessor being incorrect. The first specification only includes year-round fixed effects and yields a coefficient of -0.027 that is not statistically different than zero. The second specification includes state fixed effects, which generates a statistically significant coefficient of -0.040. Specification three adds a control for the speller's position in a round and specification four includes controls for speller's characteristics.

Moving across the columns, the estimated coefficient on "speller *i*-1 correct" moves in the expected directions. It becomes more negative when state fixed effects are added, and is unaffected by the further addition of position within round. The estimate is also stable to the further inclusion of the full battery of student characteristics, which supports the assumption that ability is uncorrelated across subsequent spellers, conditional on the controls.

These results imply that when the predecessor is correct, a speller is 4 percentage points more likely to make a mistake, relative to the predecessor being incorrect. The inferences based on marginal effects from probit models (evaluated at the mean) are the same.[12] Taking the unconditional error rate of 29.3 percent, this is a 13.3 percent increase in the probability of making a mistake.

5.2 Instrumental Variables

The results from the instrumental variables specifications are reported in Table 4. The first column shows the results for standard IV from the overidentified model, with the focal second stage estimates shown in the top panel and the first stage estimates shown in the bottom panel. The IV estimate of the coefficient on "speller *i*-1 correct" is statically significant and equal to -0.188. The p-value on the F-statistic for the excluded instruments confirms that the instruments are highly relevant. Further, the first stage estimates reveal directionally intuitive relationships. That is, one would expect experience and age to positively affect performance and homeschoolers have historically done very well at The Bee. Finally, the hypothesis that the instruments are valid is not rejected by the overidentification test.

The next column presents an analogous set of estimates but uses the performance of speller *i*+1. The second stage coefficient equals 0.049 and is not statistically different than zero. Since the predicted performance of the person spelling after you should not affect your performance,

[12] Presented in Appendix Table 2.

finding no effect is intuitive and does not ring any alarm bells about the instruments or identification.

The next two columns display results from a just identified model with speller *i*-1's (*i*+1's) experience as the instrument. Results are very similar to the overidentified model. The last two columns present estimates from the GKN Method and again, results are relatively unchanged.

Further supporting these approaches, speller *i*-1's predicted performance does not help to explain own predicted performance, in regressions that also control for state and year-round fixed effects and position within a round.[13]

Overall, Table 4's estimates of the effect of speller *i*-1's performance are all much larger than the OLS estimate. The overidentified model's estimate of 18.8 percentage points and the unconditional error rate of 29.3 percent translate into a 64 percent increased probability of an error when the predecessor is correct, relative to when the predecessor is incorrect. This compares to the OLS estimates of 13.3 percent, suggesting the OLS results are biased toward zero due to correlated shocks or that spellers respond less to peer success that is due to luck.

6. Extensions

6.1 Five Predecessors

Table 5 displays OLS and IV results when including the five previous spellers' performances.[14] In the OLS specification, the coefficients on both speller *i*-1 and speller *i*-2 are negative and significantly different than zero. After the second speller, the coefficients on the

[13] Results of this test of random assignment are presented in Appendix Table 3, column 3.

[14] The instrumental variables specification has five endogenous variables. Instruments include student characteristics for each of the five preceding spellers.

other spellers' performances are small in magnitude and not statistically significant. A similar pattern appears in the IV specification, except, as before, the coefficients are greater in magnitude than OLS coefficients, and the coefficient on speller *i*-2 is just out of standard levels of statistical significance.[15] Overall, this gives some indication that the first few predecessors are impacting a speller's performance but the less proximate predecessors are not.

6.2 Effects by Round

Table 6 explores whether the results differ by round of competition. The results for both OLS and the overidentified IV models are shown first for all rounds, and then estimated separately for preliminaries, quarterfinals, and semifinals. The OLS coefficient estimates for a predecessor being correct are statistically significantly negative for the early rounds, but not for the semifinals. The point estimates for the early rounds are also more than twice as large as for the semifinals. The IV estimates show a similar pattern across the columns, though the estimate is only statistically significant for the preliminaries.[16]

In all, the results suggest that a speller has a greater probability of making an error when her immediate predecessor is correct in the first two rounds, but perhaps is not affected in later rounds. To the extent there are differences across rounds, these could arise from changes in the

[15] In the first stage regressions, all p-values on the F-statistics for the excluded instruments are below 0.00. The overidentified model fails to pass the overidentification test. The just identified model estimates the coefficient on "speller *i*-1 correct" to be -0.132 (0.122), which is not statistically different than the overidentified model's estimate. All other coefficients are not statistically different than zero.

[16] In the just identified IV model, the preliminary round estimate is -0.39 (0.19). Quarterfinal and semifinal round estimates do not yield results that are statistically different than zero.

conditions of competition (e.g., fatigue or increased intensity) or the composition of spellers (e.g., changes in ability or experience). The next section explores whether composition could play a role, by testing for differences in responsiveness by speller characteristics.

6.3 Effects by Individual Characteristics

Table 7 displays OLS and IV results of tests for differential effects by gender and experience. The specifications include the interaction of whether speller *i*-1 is correct with an indicator for the speller's gender (equals one if male) or experience.[17] As Table 7 demonstrates, the coefficient on the gender interaction term is not statistically different than zero. The non-interaction term also fall out of significance, in part because of an increase in the standard errors' magnitudes. The coefficient on the experience interaction term is positive and marginally significant, which indicates that more experience can mitigate the negative effects of a predecessor being correct. For example, when a speller has one year of experience and her predecessor is correct, she is 25 percentage points more likely to make an error than if her predecessor is incorrect. Comparatively, a speller with two years of experience is only 11 percentage points more likely to make an error than when her predecessor is incorrect.

7. Discussion

This research shows evidence that spellers' performances are affected by their immediate predecessors and less so by less proximate peers. It is difficult to know the exact reason spellers

[17] Instrumental variables specifications have two endogenous variables. Results presented are for the overidentified model and the p-values on the F-statistics of the instruments are below 0.00. In the just identified model, estimates on the interaction term of "speller *i*-1 correct" with male and experience are 0.063 (0.272) and 0.140 (0.114), respectively.

are affected by one another but most involve some degree of psychology. In a perfectly rational model, students are endowed with an ability level and either do or do not know the word. If they don't, they put forth their best effort and guess. But results indicate that there exist factors, other than innate ability, causing a speller's performance to be affected by others' performances.

One potential explanation is an increase in pressure that has been documented to negatively affect performance in both the social psychology literature (e.g. Baumeister 1984; Butler and Baumeister 1998; Kleine et al. 1988) and the economics literature (e.g. Ariely et al. 2009; Dohman 2008; Paserman 2011). If the immediate predecessor is correct, it becomes important to be correct and consequently there is an increase feeling of pressure, relative to peers spelling incorrectly. Moreover, perhaps the speller simply uses the predecessor as a salient proxy for how all other spellers perform. That is, since the immediate predecessor was correct, the speller believes many predecessors have been correct, which may also increase the pressure.

Alternatively, confidence or optimism is a potential mechanism. Psychology and economic research has found that underconfidence can be self-fulfilling (Battle 1965; Feather 1966) and that overconfidence can have negative consequences (Barber and Odean 2001). In The Bee, a predecessor being correct may cause a speller to believe the competition is easy. Consequently, the speller is overconfident and puts in less effort, leading to a higher probability of a mistake.

These are only two potential mechanisms among many possible mechanisms. Whatever the mechanisms, spellers are somehow responding to their peers' performances. Hopefully future research will uncover or create settings with similar features that would allow the specific mechanisms at play to be identified.

Finally, there is no evidence that genders are differentially affected by predecessors. This is interesting given the existence of females differences in other competitive environments (Gneezy

et al. 2003; Gneezy and Rustichini 2004; Niederle and Vesterlund 2007; Paserman 2011). However, it should be noted that students do not randomly enter The Bee and a lack of a gender difference may be driven by selection.

References

- Angrist, Joshua D. and Kevin Lang (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *The American Economic Review.* Vol. 94, No. 5, pp. 1613-1634.
- Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar (2009): "Large Stakes and Big Mistakes," *Review of Economic Studies*. Vol. 76, Issue 2, pp. 451-469.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2005). "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, Vol. 120, No. 3, pp. 917-962.
- Barber, Brad and Terrance Odean (2001). "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics*. 141:261–292.
- Battle, E.S. (1965). "Motivational Determinants of Academic Task Persistence," *Journal of Personality and Social Psychology*. Vol. 2, 209-218.
- Baumeister, Roy F (1984). "Choking under pressure: Self-consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology.* Vol. 46(3).
- Brown, Jennifer (2011). "Quitters Never Win: The (Adverse) Incentive Effects of Competing With Superstars," *Journal of Political Economy*. Vol. 119, No. 5. pp. 982-1013.
- Bull, Clive, Andrew Schotter, and Keith Weigelt (1987). Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*. Vol. 95, No. 1. pp. 1-33.

- Butler, J. L., and R. F. Baumeister (1998). "The Trouble With Friendly Faces: Skilled Performance With a Supportive Audience," *Journal of Personality and Social Psychology*. 75(5), 1213–1230.
- Dohman, Thomas (2008). "Do Professionals Choke Under Pressure?" *Journal of Economic Behavior & Organization*. Vol. 65, Issue 3-4, pp. 636-653.
- Feather, N.T. (1966). "Effect of Prior Success and Failure on Expectations of Success and Subsequent Performance," *Journal of Personality and Social Psychology*. Vol. 3, 287-298.
- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini (2003). "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, CXVIII, pp. 1049--1074.
- Gneezy, Uri, and Aldo Rustichini (2004). "Gender and Competition at a Young Age," *American Economic Review Papers and Proceedings*. XCIV, pp. 377--381.
- Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo (2009). "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*. Vol. 1(4).
- Hoxby, Caroline (2000). "Peer Effects in the Classroom: Learning From Gender and Race Variation," *NBER Working Paper #7867*.
- Kleine, D., Sampedro, R. and Lopes Melo, S. (1988). "Anxiety and performance in runners. – Effects of stress on anxiety and performance in ergometric tests and competitions," *Anxiety Research*. Vol. 1, pp. 235-246.

- Lavy, Victor, Olmo Silva, and Felix Weinhardt (2009). "The Good, the Bad and the Average: Evidence on the Scale and Nature of Ability Peer Effects in Schools, *NBER Working Paper #15600.*
- Lazear, Edward P., and Rosen, Sherwin (1981). "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*. Vol. 89, pp. 841-64.
- Lazear, Edward (2000). "Performance Pay and Productivity," *American Economic Review.* Vol. 90, No. 5. pp. 1346-1361.
- Mas, Alex and Enrico Moretti (2009). "Peers at Work," *American Economic Review.* Vol. 99(1), pp. 112-45.
- Nalebuff, Barry J., and Stiglitz, Joseph E (1983). "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell J. Econ*. 14: 2 pp. 1-43.
- Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*. pp. 1067-1101.
- Paserman, M. Daniele (2011). "Gender Differences in Performance in Competitive Environments: Field Evidence from Professional Tennis Players," Working Paper.

Table 1: Error Rate by Round

Round*	Round Name**	Observations	Error Rate
1st	Preliminaries	1,386	27.3%
2nd	Quarterfinals	485	20.8%
3rd	Semifinals	384	36.2%
4th	Semifinals	245	39.6%
5th	Semifinals	148	36.5%
6th	Semifinals	94	36.2%
7th	Semifinals	60	23.3%
8th	Semifinals	46	28.3%
9th	Semifinals	28	35.7%
10th	Semifinals	18	44.4%
11th	Semifinals	4	0.0%
12th	Semifinals	4	0.0%
13th	Semifinals	4	25.0%
All	All	2,906	29.3%

*The round is the on-stage round. Written test rounds are excluded in the ordinal count.

**Round name is defined by the 2008 competition. Championships final rounds, which are excluded, occur when two or three spellers remain.

Table 2: Summary Statistics

Observation level = speller

Variable	Obs	Mean	Std. Dev.	Min	Max
Male	1,386	0.49	0.50	0	1
Experience*	1,386	1.27	0.61	1	5
Age	1,386	12.78	1.15	8	15
Grade	1,386	7.21	1.02	2	8
Black	1,386	0.04	0.19	0	1
Hispanic	1,386	0.03	0.18	0	1
South Asian	1,386	0.11	0.32	0	1
Asian	1,386	0.11	0.32	0	1
Other	1,386	0.04	0.19	0	1
Homeschool	1,386	0.12	0.33	0	1

Observation level = speller-round

	Obs	Mean	Std. Dev.	Min	Max
Speller i-1 correct	2,857	0.71	0.45	0	1
Position in round (percentile)	2,906	0.51	0.29	0	1
Preliminary indicator (Round 1)	2,906	0.48	0.50	0	1
Quarterfinal indicator (Round 2)	2,906	0.17	0.37	0	1
Semifinal indicator (Rounds > 2)	2,906	0.36	0.48	0	1

*Experience is equal to the number of years the speller has been in The Bee (including current year).

Table 3: Effect of Immediate Predecessor's Performance

Linear Probability Models, Dependent Variable = 1 if Spelled Correctly

Variable	(1)	(2)	(3)	(4)
Speller i-1 correct	-0.027	-0.040**	-0.040**	-0.039**
	(0.019)	(0.020)	(0.020)	(0.019)
Position in round (percentile)	--	--	-0.143	-0.094
	--	--	(0.162)	(0.166)
State Fixed Effects	No	Yes	Yes	Yes
Controls for Student Characteristics	No	No	No	Yes
Observations	2,857	2,857	2,857	2,857
R-squared	0.062	0.096	0.096	0.134

Notes: An observation is a speller in a round using data from the 2004-2008 competitions. First speller in each round is excluded. Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include year-round fixed effects. Student characteristics include experience, gender, age, grade, race/ethnicity, and a homeschool dummy.

Table 4: Effect of Immediate Predecessor's Ability
Instrumental Variables and GKN Method
Instruments = Speller i-1 or Speller i+1 Characteristics

Instrumented Variable	Instrumental Variables[1]				GKN Method[2]	
	Speller i-1	Speller i+1	Speller i-1	Speller i+1	Speller i-1	Speller i+1
Speller i-1 (i+1) correct	-0.188**	0.049	-0.194*	0.005	--	--
	(0.088)	(0.094)	(0.116)	(0.121)	--	--
Speller i-1 (i+1) predicted performance	--	--	--	--	-0.188***	-0.011
	--	--	--	--	(0.065)	(0.065)
Observations	2,857	2,857	2,857	2,857	2,857	2,857
First-Stage Results						
Instruments						
Speller i-1 (i+1) experience	0.082***	0.076***	0.094***	0.090***	--	--
	(0.011)	(0.011)	(0.011)	(0.011)	--	--
Speller i-1 (i+1) gender	0.049***	0.044***	--	--	--	--
	(0.017)	(0.017)	--	--	--	--
Speller i-1 (i+1) grade	0.047***	0.041***	--	--	--	--
	(0.018)	(0.018)	--	--	--	--
Speller i-1 (i+1) homeschooled	0.068***	0.058***	--	--	--	--
	(0.026)	(0.027)	--	--	--	--
P-value of F-statistic on instruments	0.000	0.000	0.000	0.000	--	--
Sargan's Overidentification Test p-value	0.256	0.521	--	--	--	--

Notes: An observation is a speller in a round using data from the 2004-2008 competitions. First (last) speller in each round is excluded when using speller i-1 (i+1). Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include controls for position of speller, state fixed effects, and year-round fixed effects as well as controls for experience, gender, age, grade, race/ethnicity, and a homeschool dummy.

1. Other excluded instruments (not shown) in the overidentified model include the speller i-1's (i+1's) age and dummies for race/ethnicity.

2. GKN Method uses predicted performance of speller i-1(or i+1). Predicted performance is based on the speller's gender, age, grade, race/ethnicity, a homeschool dummy, experience, state, and round. Prediction based on all spellers from all years other than current year.

Table 5: Effect of Predecessors' Performances

OLS and IV, Dependent Variable = 1 if Spelled Correctly

Variable	OLS	IV[1]
Speller i-1 correct	-0.038*	-0.178**
	(0.020)	(0.076)
Speller i-2 correct	-0.053***	-0.100
	(0.020)	(0.078)
Speller i-3 correct	-0.024	-0.043
	(0.020)	(0.078)
Speller i-4 correct	-0.012	-0.072
	(0.020)	(0.077)
Speller i-5 correct	0.012	-0.051
	(0.021)	(0.078)
Observations	2664	2664
R-squared	0.139	--
Sargan's Overidentification Test p-value	--	0.022

Notes: An observation is a speller in a round using data from the 2004-2008 competitions. The first five spellers in each round are excluded. Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include controls for position of speller, state fixed effects, and year-round fixed effects as well as controls for experience, gender, age, grade, race/ethnicity, and a homeschool dummy.

1. Excluded instruments include speller i-1's, i-2's, i-3's, i-4's i-5's experiences, genders, ages, grades, races/ethnicities, and whether homeschooled.

Table 6: Effect of Immediate Predecessor's Performance by Rounds

OLS and IV, Dependent Variable = 1 if Spelled Correctly

	OLS			
Variable	All Rounds	Preliminaries (Round 1)	Quarterfinals (Round 2)	Semifinals (Rounds > 2)
Speller i-1 correct	-0.039**	-0.066**	-0.096*	-0.036
	(0.019)	(0.027)	(0.049)	(0.033)
Observations	2857	1381	480	996
R-squared	0.134	0.174	0.200	0.171
	Instrumental Variables [1]			
Variable	All Rounds	Preliminaries (Round 1)	Quarterfinals (Round 2)	Semifinals (Rounds > 2)
Speller i-1 correct	-0.188**	-0.255**	-0.319	-0.136
	(0.088)	(0.106)	(0.202)	(0.130)
Observations	2857	1381	480	996
P-value of F-statistic on instruments	0.000	0.000	0.020	0.000
Sargan's Overidentification Test p-value	0.256	0.688	0.329	0.243

Notes: An observation is a speller in a round using data from the 2004-2008 competitions. First speller in each round is excluded. Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include controls for position in round, state fixed effects, and year-round fixed effects and experience, gender, age, grade, race/ethnicity, and a homeschool dummy.

1. Excluded instruments include speller i-1's experience, gender, age, grade, race/ethnicity, and whether homeschooled.

Table 7: Effect of Immediate Predecessor's Performance
By Individual Characteristics
OLS and IV, Dependent Variable = 1 if Spelled Correctly

Variable	OLS		IV[1]	
Speller i-1 correct	-0.036	-0.059	-0.180	-0.391**
	(0.028)	(0.037)	(0.138)	(0.163)
Speller i-1 correct x Male	-0.005	--	-0.024	--
	(0.038)	--	(0.194)	--
Speller i-1 correct x Experience	--	0.013	--	0.140*
	--	(0.020)	--	(0.083)
Observations	2,857	2,857	2,857	2,857
R-squared	0.134	0.134	--	--
Sargan's Overidentification Test p-value	--	--	0.0925	0.243

Notes: An observation is a speller in a round using data from the 2004-2008 competitions. First speller in each round is excluded. Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include controls for position in round, state fixed effects, and year-round fixed effects, as well as experience, gender, age, grade, race/ethnicity, and a homeschool dummy.

1. Excluded instruments include speller i-1's experience, gender, age, grade, race/ethnicity, and whether homeschooled and their interactions with speller i's gender (or experience).

Peers, Pressure, and Performance at the National Spelling Bee

Web Appendix

A.1 The Bee from 1997-2003

Prior to 2004, The Bee had two notable changes. First, between 1997 and 2001, there were no preliminary rounds and so similar to later rounds, an incorrect spelling necessarily results in elimination. Second, in the 2002 and 2003 preliminary rounds, the oral round came before the written test, which was eventually reversed in 2004.

This Appendix discusses results when including competitions between 1997 and 2003.[1] This allows for more precise inference and to test whether results hold in past years. However, when pooling the data, the early rounds differ and this may account for slight differences in the results. In addition, the demographic controls are only available between 2004 and 2008, which is another reason the main text focuses on the later years.

A.1.1 Additional Data

Determined entirely by name, I construct a gender variable for spellers in the 1997-2003 competitions. However, gender neutral names and foreign names not familiar to the author are difficult to determine. With the help of myspace.com, the gender is assigned based on the highest proportioned gender as determined by the website, which generally solves the foreign name issue.[2] As a final check on gender, I similarly assigned gender to the 2005 data. I then

---

[1] Experience is missing in 1996 so this year is dropped. Excluding 1997 or 1998, the years that may inaccurately measure experience for very few spellers, does not change the results.

[2] To check the validity of gender from 1997-2003, a second person performed the task of assigning gender to names. The assignments match one another for 98.5 percent of the 1,740 spellers. All results hold when using the second person's assignments.

compared my assignments to their true gender from their biographies and had a match rate of over 95 percent.[3]

Appendix Table 1 displays summary statistics for the full sample (1997-2008) and the early years (1997-2003). The statistics on the variables are comparable to the sample in the main text.

A.2 Effect of Immediate Predecessor Using 1997-2003 Data

Appendix Table 2 estimates the effect of the immediate predecessor being correct on a speller's performance while using data from 1997-2003. The first set of data uses 1997-2008, the second set only uses 1997-2003, while the third set uses 2004-2008, as in the text.

The estimated effects in the 1997-2003 data are never statistically different than zero, whether it be OLS, probit, or IV. Consequently the full sample, 1997-2008 has muted estimates compared to the 2004-2008 sample. It is difficult to determine if these muted effects are because the spellers are substantially different than the spellers in the 2004-2008 competitions or if the change in the structure of the competition had an impact.

I also run a test of random assignment of speller order. I regress speller *i*'s predicted performance, using the GKN Method, on speller *i*-1's predicted performance, with a full set of controls. Appendix Table 3 displays these results.

When using the full sample, with a limited number of controls, the coefficient on speller *i*-1's predicted performance is marginally statistically different than zero. However, neither of the subsamples can reject that the coefficient is different than zero or that assignment is not random.

[3] All results hold when excluding observations with foreign names unknown to the author and gender neutral names.

Appendix Table 1: Summary Statistics
Using Years 1997-2003

	Observation level = speller									
	1997-2008					1997-2003				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Male	3,126	0.48	0.50	0	1	1,740	0.47	0.50	0	1
Experience*	3,126	1.25	0.55	1	5	1,740	1.22	0.50	1	4
	Observation level = speller-round									
Speller i-1 correct	7,743	0.70	0.46	0	1	4,886	0.69	0.46	0	1
Position in round (percentile)	7,858	0.51	0.29	0	1	4,952	0.51	0.29	0	1
Preliminary indicator (Round 1)	7,858	0.40	0.49	0	1	4,952	0.35	0.48	0	1
Quarterfinal indicator (Round 2)	7,858	0.20	0.40	0	1	4,952	0.22	0.42	0	1
Semifinal indicator (Rounds > 2)	7,858	0.40	0.49	0	1	4,952	0.42	0.49	0	1

*Experience is equal to the number of years the speller has been in The Bee (including current year).

Appendix Table 2: Effect of Immediate Predecessor's Performance

OLS, Probit, IV, Dependent Variable = 1 if Spelled Correctly

Using Years 1997-2003

Variable	OLS			Probit[1]			IV[2]		
	1997-2008	1997-2003	2004-2008	1997-2008	1997-2003	2004-2008	1997-2008	1997-2003	2004-2008
Speller i-1 correct	-0.021*	-0.014	-0.039**	-0.023*	-0.017	-0.039**	-0.076	0.047	-0.188**
	(0.012)	(0.015)	(0.019)	(0.012)	(0.015)	(0.019)	(0.072)	(0.093)	(0.088)
Full Set of Controls for Student Characteristics	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7,743	4,886	2,857	7,715	4,876	2,838	7,743	4,886	2,857
R-squared	0.117	0.123	0.134	--	--	--	--	--	--
P-value of F-statistic on instruments	--	--	--	--	--	--	0.000	0.000	0.000
Sargan's Overidentification Test p-value	--	--	--	--	--	--	0.131	0.125	0.256

parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions include controls for position in round, state fixed effects, and year-round fixed effects, as well as experience and gender. When using only the 2004-2008 sample, there are also controls for age, grade, race/ethnicity, and whether homeschooled.

1. Marginal effects evaluated at the mean are reported.

2. Excluded instruments include speller i-1's experience and gender. When using only the 2004-2008 sample, instruments also include speller i-1's age, grade, race/ethnicity, and whether homeschooled.

Appendix Table 3: Testing Random Assignment

Linear Probability Models, Dependent Variable = Speller i's Predicted Performance [1]

Variable	1997-2008	1997-2003	2004-2008
Speller i-1 predicted performance[1]	-0.017*	0.018	-0.020
	(0.010)	(0.012)	(0.017)
Observations	7,743	4,886	2,857
R-squared	0.730	0.797	0.539

Notes: An observation is a speller in a round using data from the 1997-2008 competitions. First speller in each round is excluded. Robust standard errors in parantheses. *** means significant at 1% level, ** at 5%, and * at 10%. All regressions control for state fixed effects and year-round fixed effects.

1. Uses GKN method. Performance is predicted as the linear probability of speller being correct based on the speller's gender, experience, position in round, state, and round. When using the 2004-2008 sample, it also includes age, grade, race/ethnicity, and a homeschool dummy. Prediction based on all spellers from all years (in the sample) other than current year.