

# Does Speaking Order Matter in Individual Events Competition?

JEROLD L. HALE and  
FRANKLIN J. BOSTER\*

The belief by forensic mentors and competitors that speaking order influences competitive outcomes is not new. Concern about order effects was significant enough to prompt Franklin Knower to investigate the issue four decades ago, and that concern still exists today.<sup>1</sup> Tournament lore is replete with stories of competitors, sometimes at the urging of a coach, arriving late to a round of competition. While some lack of punctuality is due to double and triple entries, some competitors try to improve their speaking position and, hence, the critic's evaluation. But does speaking order really influence competitive outcomes?

To answer this question we will review the available research findings on order effects in forensic competition. We know of four studies that address the issue of order effects. Those studies have produced seemingly inconsistent and contradictory findings. A closer examination of the research, however, indicates that the findings are consistent, and that the impact of speaking order is weak.

Our discussion of the body of research will progress chronologically. A brief discussion of each study will be presented and followed by a summary of the research findings.

## *Knower, 1940*

The data for Knower's study of order effects were ranks assigned to competitors in the National tournaments of the National Forensic League, Phi Rho Pi, the Northern Oratorical League, the Intercollegiate Oratorical Association, and Pi Kappa Delta. He made a total of 13,265 observations. Knower reported pervasive order effects and offered several specific ranks. Finally, he posited that fourth, fifth, and sixth speakers were more likely to be assigned a rank of first than other speakers.<sup>2</sup>

---

\*The National Forensic Journal, IV (Spring 1986), pp. 45-51.

JEROLD L. HALE is Assistant Professor of Communication, Miami University, Oxford OH 45056 and FRANKLIN J. BOSTER is Associate Professor of Communication, Arizona State University, Tempe 85281.

<sup>1</sup>Franklin H. Knower, "A Study of Rank Order Methods of Evaluating Performances in Speech Contests," *Journal of Applied Psychology* 24 (1940), pp. 633-644.

<sup>2</sup>Knower, pp. 633-644.

While Knower's findings reinforced the beliefs of much of the forensic community they should be viewed with a healthy amount of skepticism. Only the frequency with which a speaking position was assigned a given rank was reported. No tests of statistical significance or measures of association were performed. Put differently, it is not possible to determine from the way the data were reported whether order effects occur with greater than chance regularity or the strength of the relationship between speaking order and rank assignments outcomes. Furthermore, the finding that extreme speaking positions are more likely to receive an intermediate rank than an extreme rank is a function of mathematical probability and not necessarily evidence of an order effect. An example will clarify this argument.

For any panel of five or more speakers the probability of being assigned an intermediate rank is greater than the probability of being assigned an extreme rank. Assume, for a moment, that the probability of being assigned any rank is equal. In a group of five speakers the probability of being assigned any rank would be .20, i.e., it would happen one-fifth of the time. The probability of being assigned an intermediate rank would be .20 multiplied by the number of intermediate ranks, in this case three. The probability of being assigned an extreme rank would be .20 multiplied by the number of extreme ranks, or two. If order had no impact on competitive outcomes the probability of receiving an intermediate rank would be .60, i.e., it would happen six times out of ten. The probability of being assigned an extreme rank would be .40, or four times out of ten.<sup>3</sup> The point to be made is that Knower's finding that first and last speakers were more likely to receive intermediate ranks than extreme ranks is expected and cannot, by itself, be evidence of an order effect.<sup>4</sup>

*Becker, 1953*

Samuel Becker attempted to provide more definitive conclusions about the impact of speaking order. He examined the relationship between speaking order and ranks assigned in 22 years of Northern Oratorical League competition. His research included a total of 660

---

<sup>3</sup>This argument assumes that speakers were randomly assigned to speaking positions. Random assignment of speakers would mean that speaker quality should be independent of speaking position. For most tournaments this is a reasonable assumption.

<sup>4</sup>James A. Benson and Susan K. Maitlen, "An Investigation of the Relationship Between Speaking Order and Rank Assignment in Forensic Competition," *Journal of the American Forensic Association* 11 (1975), 183-188; Jerold L. Hale, "The Effect of Speaking Order on Rank Assignments and Quality Ratings for Extemporaneous Speeches," Speech Communication Association Convention, Anaheim, November, 1981.

observations. For each speaking position Chi-square tests were performed. The Chi-square test is a statistical significance test that indicates whether the observed rank assignments differ from expected rank assignments more than would be expected by chance.

From those data Becker argued that speaking order influenced the ranks assigned to the first three speakers but not to the latter three in panels of six. More specifically, he argued the first speaker was more likely to be assigned a rank of fourth than other ranks. The second speaker was more likely to be assigned a rank of second than other ranks, and the third speaker was more likely to be assigned a rank of fifth than other ranks. Becker's findings appear to be similar to Knower's, though the extent of the similarity is difficult to discern because of the manner in which Knower's data were reported.<sup>5</sup>

*Benson and Maitlen, 1975*

James Benson and Susan Maitlen investigated the effect of speaking order on rank assignments in three individual events tournaments. A total of 584 observations were made. For each speaking position Chi-square tests and Kolomogorov-Smirnov tests were performed. The Kolomogorov-Smirnov test, like the Chi-square test, is a test of statistical significance. It, too, determines whether or not the ranks assigned to each speaking position differ significantly from expected rank assignments. Benson and Maitlen also did separate analyses for prepared versus non-prepared events, preliminary rounds of competition, semi-final and final rounds of competition, and for different size speaking panels.

They reported speaking order had no significant impact on intermediate rank assignments, no significant impact on being ranked first or last, and no significant impact when comparing ranks assigned in prepared versus non-prepared events. The only statistically significant order effect was found when comparing preliminary rounds of competition to out rounds. One speaking position received ranks of second more often than would be expected by chance. No other order effects were found.<sup>6</sup>

*Hale, 1982*

Jerold Hale conducted the most recent investigation of order effects in forensic competition. His approach differed from the approach used in previous studies. Prior studies used data from live tournament competition. The investigation by Hale was a laboratory experiment which simulated an extemporaneous speaking

---

<sup>5</sup>Samuel L. Becker, "The Ordinal Position Effect," *Quarterly Journal of Speech* 39 (1953); 217-219; Knower, pp. 633-644.

<sup>6</sup>Benson and Maitlen, pp. 183-188.

contest. Six speakers gave speeches ranging in length from five to seven minutes. The speeches were videotaped. The videotapes were spliced together so that each speaker was shown in each possible speaking order. The speeches were then shown to other college students who ranked each speaker and assigned a quality rating. A total of 1,044 observations were made. Hale reported a statistically significant, but extremely weak, relationship between speaking order and both ranks and ratings. For both ranks and ratings the relationship was linear. More specifically, later speakers received rankings and ratings that improved, but only very slightly.<sup>7</sup>

#### Summary of Findings

Of the four studies conducted two argued that speaking order had a pervasive impact on the outcomes of forensic competition and two argued that the impact of speaking order was trivial. While these findings appear contradictory they are consistent and indicate the relationship between speaking order and success is weak. The seeming contradiction occurs because most of the research used statistical significance tests and not measures of association. Significance tests and measures of association provide different information, and one should not be used to the exclusion of the other when conducting quantitative research.<sup>8</sup> To clearly illustrate this a distinction between statistical significance tests and measures of association must be explicated.

Statistical significance tests determine whether a relationship occurs with a certain regularity. That is, significance tests indicate the likelihood of the research results being the result of chance. For example, Becker found that speaking order had a statistically significant impact on the assignment of some ranks. Statistical significance then, is another way of saying that the likelihood of findings being a chance occurrence is small.<sup>9</sup>

Measures of association determine the strength or magnitude of a relationship between two or more variables. It is not uncommon in forensics to hear coaches or students say that two things are correlated, i.e., the two things are associated or related. For example, coaches often times attempt to convince students that effort and performance are correlated, so that the more effort the

---

<sup>7</sup>Hale.

<sup>8</sup>Harris M. Cooper, "On the Significance of Effects and the Effects of Significance," *Journal of Personality and Social Psychology* 41 (1981), pp. 1013-1018.

<sup>9</sup>Becker, pp. 217-219. The conclusions we draw about the differences between significance tests and strength of effect measures are found in several statistical publications. However, the Cooper citation summarizes the issue nicely.

student puts forth the better he or she will do in competition. The most common measure of association is the Pearson Product Moment Correlation, or Pearson's  $r$ .<sup>10</sup> The correlation coefficient,  $r$ , measures the strength of a relationship. Pearson's  $r$  can range in numerical value from -1.00 to 1.00. If effort and performance were correlated -1.00 they would be perfectly negatively correlated. As effort increased performance quality would decrease. If the correlation between effort and performance is zero effort and performance quality would be unrelated. Finally, if effort and performance were correlated 1.00 the two would be perfectly and positively related. As effort increased performance quality would increase.

Cast in terms related to speaking order, a positive correlation between speaking order and rank assignments would indicate that later speakers would receive higher rank assignments. A negative correlation between speaking order and rank assignments would mean that later speakers received lower rank assignments. The more a correlation coefficient deviates from zero the stronger the relationship between two variables would be.

A statistically significant relationship does not mean that a relationship is strong, only that it is observed with greater than chance regularity. Statistical significance, is, in part, a function of the sample size or number of observations made. As the sample size increases statistical tests become more powerful, or more likely to detect a subtle effect if one exists. For example, in Becker's research 660 observations were made. With a sample of 660 a correlation of .08 or larger would be statistically significant. Typically, correlations of .20 or less are considered to be small effect sizes, or indicative of weak relationships.<sup>11</sup>

How does this discussion of the difference between statistical significance and measures of association demonstrate the consistency among diverse research findings? Two of the four studies reviewed found statistically significant order effects. While Knower did not perform statistical significance tests he reported pervasive order effects. On the other hand, three of the four studies conducted found extremely weak relationships between speaking order and competitive success. Only Knower's findings are not indicative of a weak relationship and he did not test for the strength of the relationship. Furthermore, the data are not reported in such a way to reconstruct the data set or to perform the tests in retrospect.<sup>12</sup>

---

<sup>10</sup>Karl Pearson, "On the General Theory of the Influence of Selection on Correlation and Variation," *Biometrika* 8 (1911-1912), 437-443.

<sup>11</sup>James Cohen, *Statistical Power Analysis for the Behavioral Sciences* (New York: Academic Press, 1977).

<sup>12</sup>Knower, pp. 633-644.

A discussion of the effect sizes found in the other three studies would make our argument more compelling. Becker did not perform measures of association when analyzing those data.<sup>13</sup> He did, however, report the data with enough detail so that subsequent analyses could be conducted. When we performed those analyses the strongest correlation between speaking order and performance was  $r = .15$ . In short, Becker observed a statistically significant relationship between speaking order and competitive success but speaking order had a trivial impact on rank assignments.<sup>14</sup>

Benson and Maitlen did not perform measures of association in analyzing those data. It is reasonable to infer, however, that the relationship between speaking order and rank assignments was weak. While a statistically significant relationship is not an indication of a strong relationship, an insignificant relationship is indicative of a weak relationship, especially when the sample size is large. The sample size in that research was 550. The effect size in the study could not have exceeded  $\pm .08$  without producing statistically effects.<sup>15</sup>

Hale did perform measures of association. He used a measure of association called Eta squared which can be transformed easily to correlation coefficients. Speaking order and rank assignments were correlated  $r = .12$ , and speaking order and quality ratings were correlated  $r = .08$ . While some might criticize the method employed in Hale's research, e.g., because it included videotaped speeches and used college students as critics, the latter of the two criticisms seems to be a further indication that speaking order has little impact. If college students with no forensic training were uninfluenced by speaking order the likelihood of skilled coaches and judges being influenced by speaking order would be even smaller.<sup>16</sup>

The available studies do indicate that speaking order has a statistically observable impact on rank assignments and quality ratings. That finding, considered alone, is deceptive. Statistical significance tests are influenced by the sample size used in the research. As the sample size increases a significance test is more likely to detect weak relationships.<sup>17</sup> Using significance tests without calculating the strength of the relationship between two

---

<sup>13</sup>Becker, pp. 217-219. We are critical of the statistical analyses used by Professors Knower and Becker. In fairness to them we should point out that measures of association received more widespread use as computer programs and algorithms simplified their calculations.

<sup>14</sup>Becker, pp. 217-219; A complete copy of the analyses available upon request from the first author.

<sup>15</sup>Benson and Maitlen, pp. 183-188.

<sup>16</sup>Hale.

<sup>17</sup>Cohen.

variables can produce misleading conclusions. The literature on order effects in forensic competition is a good example. When the strength of the relationship between speaking order and success measures is calculated the data emphatically indicate that speaking order has a trivial impact on competitive outcomes.

Some readers might be tempted, based on these results, to conclude that attempting to improve speaking position does not diminish one's chances of success and on rare occasions it could improve one's chances. That conclusion would be short sighted. Three studies report which speaking position received the best rank assignments. For each of the three studies a different speaking position is slightly better.<sup>18</sup> Contestants trying to improve their competitive standing by some trivial degree would merely be guessing as to an effective speaking order strategy.

A more realistic interpretation of the available research would be that time spent either worrying about one's speaking position or trying to improve positions is time wasted. That time would be more productively spent in preparation.

---

<sup>18</sup>Becker, pp. 217-219; Benson and Maitlen, pp. 183-188; Hale; Knower, pp. 633-644.