INEXPERIENCED AND EXPERIENCED DEBATE JUDGES: BEYOND "NAME CALLING"[1]

E. Sam Cox
Central Missouri State University

Tammy Lyman Honse
Central Missouri State University

Forensic scholars have devoted considerable attention to the role of the judge in academic debate. One particular focus that has caused noteworthy controversy is the use of the inexperienced or lay judge. Unfortunately, as Colbert (1989) notes, "criticism of academic competitive debate is seldom, if ever, supported by objective and carefully measured observation" (p. 1). A survey of the literature verifies that Colbert's indictment is applicable to both sides of the controversy surrounding the use of inexperienced judges.

Illustration of the subjective approaches to determining the worth of an inexperienced judge can be seen in the following. Gotcher (1985), in his article, "Who Should Judge CEDA Debate," says that the scarcity of competent judges can be documented by "reading a few debate ballots" (p. 4). Srader (1990) only asserts that lay judges are harmful to competitive debate. Likewise, Marrs (1990) insists that lay judges are capable of rendering qualified decisions, based upon the claim that "I think most judges are like me" (p. 16). Although every argument need not rely upon a carefully planned study, claims must be grounded in verifiable data if any generalizations are to be made.

Miller (1969) did conduct a study which is not based upon anecdotal information to compare inexperienced and experienced judges. However, this dated study only compares ratings, rankings and win-loss decisions but does not address reasons for decisions.

Thus, a significant gap exists in the research concerning the controversy over the inexperienced judge. The lack of recent and generalizable data prompted the following investigation. This study is not limited to a comparison of just decision rendering but is extended to include an analysis of the educational roles of both inexperienced and experienced judges.

REVIEW OF LITERATURE

Examining the judge's contribution is important because it is vital to the enhancement of the universal goals of debate: critical thinking, communication skills, and academic advancement (Lybbert, 1985). Most scholars agree that the judge's role is twofold: as a decision maker and as an educator. Richardson (1985) however, asserts that "to combine the role of debate judge and educator is a little like telling someone he must serve the dual function of defense attorney and Lord High Executioner, simultaneously" (p. 1). The rationale provided by Richardson is that a judge describes what occurred in a round while a critic/educator prescribes what should have occurred in the round. However, the prevailing opinion is that judges have the obligation to serve beyond the role of descriptor, and they are indeed capable of performing two types of evaluation simultaneously--that of a judge and that of a critic (Patterson and Zarefsky, 1983). Phillips (1985) goes so far as to posit that while both roles are important, the "debate judge is first and foremost an educator" (p.1). Herbeck (1984) concurs and argues that debate has both a competitive and educational end; Decker and Morello (1984) believe that debate is "first and foremost an educational activity" (p.154). If indeed the purpose of debate is to teach (Rowland, 1984; Ganer, 1987), judges must also serve as educators.

The National Federation of State High School Associations recognizes the educational function of debate and contends that "forensic educators must take a leadership role in promoting the teaching of oral communication" (p.55). To meet this need they passed a resolution in December, 1989, concerning the development and improvement of curriculum and instruction. The resolution states that, "Judges should provide extensive written comments on ballots, even in cases where extensive oral critiques are given" (p.56). The idea behind the resolution is that the comments on ballots need to be targeted toward enhancing the educational process of the activity.

The dominant view is that one role of the judge is to further the educational goals of debate. A pressing question is how to select qualified judges in order to insure that the activity accomplishes the purposes for which it was designed. Research by numerous people (see for example, Burgoon, 1975; Ulrich, 1984; & Cox, 1974) suggests ways in which experienced judges can enhance the educational benefits of forensics. The question has arisen, however, whether or not those outside the forensic community are also capable of insuring the fulfillment of the goals of debate.

Although there is considerable controversy regarding the use of inexperienced judges, most of the discussion is based upon anecdotal information. As early as 1940 Holm objects to the inclusion of inexperienced judges claiming that judges must be trained in the art and pedagogy of the activity (p. 5). Similarly, Floyd and Cripe (1982) posit that a competent judge is one who knows debate theory, terms and concepts. A debate critic should also be familiar with the procedures of debate, obligation of participants, and the arguments surrounding the topic, according to Balthrop (1983).

In addition to criticisms concerning the inexperienced judge's lack of knowledge regarding debate practices, others stress the danger of intervention. Srader (1990) asserts that "Lay judges tend to vote for intuitive arguments and to intervene extensively against complex arguments which seem to contradict their experience even though they may not be experts in the field" (p. 9). Colbert (1989) claims that using only inexperienced judges is as

[1].The initial design and some frequencies from this research were presented at the 1989 SCA Convention in San Francisco.

"foolhardy" as a publisher hiring inexperienced editors to make newspapers more meaningful for the average reader.

Nicolai (1985) warns of another potential harm if inexperienced judges are used. He notes how "there is a tendency to dismiss the hired judge's decision and/or comments if they seem to deviate from the norm" (p. 5). The conclusion would follow that learning is hindered if comments are ignored. Furthermore, as Colbert (1989) argues, the educational benefits are harmed when the lay judge treats each debate round as a single unit rather than "as an educational means to an ends, not an ends unto itself" (p. 2-3). Gotcher (1985) denounces the idea that anyone can judge CEDA because students are seldom given information about the philosophy of the lay judge, thus adaptation is prevented and consequently the value of debate is harmed.

A final criticism of inexperienced judges deals with the inability to record the development of arguments throughout the debate. Freeley (1986) lists one of the functions of a judge as taking notes during the debate. He argues that "Experienced educators who have judged thousands of debates are known for the care with which they take notes during a debate" (p. 280-281). Phillips (1985) concurs by directly claiming, "Judges should take flows," and that lay judges should not be used in academic debates (p. 13).

There are those, however, who believe that inclusion of inexperienced judges is beneficial to academic debate. Patterson and Zarefsky (1983) list minimum criteria any judge critic must meet. Although they are not arguing for or against the use of inexperienced judges, they establish boundaries necessary for sound academic debate. They claim that debaters can legitimately expect judges to be: interested and attentive, unbiased, intelligent, careful listeners, reasonable regarding inferences, and able to apply common sense (p. 295). Therefore, anyone who meets these criteria should be able to competently judge a debate round regardless of experience. In fact, Durkee (1985) notes that most inexperienced judges have equally sound decision making skills and are as careful in rendering decisions as the best expert judges. He goes so far as to claim that it is "distasteful to require a minimum level of knowledge about debate in order to qualify as a judge" (p. 6).

Tournament directors are urged to include inexperienced critics in the judging pool for several reasons. Brydon (1984) suggests that debate needs to be performed in front of real audiences in order to learn to adapt to persons besides trained debate critics. He argues that "University faculty in communication and other disciplines, who are not involved in debate coaching, are certainly competent judges of argumentation and persuasion" (p. 87). Similarly, Friedman (1971) posits that lay judges would enable debaters to assess any audience and be able to convince "the man in the street of the propriety of his position" (p. 126). Hollihan, Riley, and Austin (1983) describe the CEDA vision as where students relate their arguments "to trained and untrained judges, equally well" (p. 872). Even though Srader

(1990) generally opposes reliance upon lay judges, he acknowledges that their inclusion enhances training in communication skills and insures the indispensable lesson of adaptability.

Durkee (1985) and Davis (1990) deny the fact that inexperienced judges distort the educational nature of the activity. Rather, they claim lay judges serve as a bridge between the artificial environment of a debate round and the real world. Marrs (1990) concurs by stressing how inexperienced judges provide the "one point of contact with reality in a field fraught with the danger of living in an ivory tower" (p. 16).

By not including inexperienced judges, Davis (1990) stresses how students are taught contempt for persuasion as well as argumentation, which in turn feeds the harm of elitism and reduces the educational legitimacy of the activity. He believes lay judges should be scheduled as the specialist judges are now.

Although arguments have been advanced for and against the use of inexperienced judges, none has been based upon anything except anecdotal information. Miller (1969), however, conducted research to determine if there was a difference between inexperienced and experienced judges. The study reveals that when comparing coaches and non-coaches regarding wins, team ratings and speaker points, the only significant difference was in relation to speaker points. Miller's study, nevertheless, is dated and very limited in scope, i.e., it does not consider the main vehicle for education--reasons for decisions.

Therefore, due to the absence of recent verifiable data concerning inexperienced judges, the following research questions were generated.

Research Question #1: Will there be a significant difference between decisions by inexperienced and experienced judges?

Research Question #2: Will speaker ratings differ between inexperienced and experienced judges?

Research Question #3: Will reasons for decision differ between inexperienced and experienced judges? If so, how will they differ?

METHODOLOGY

At a debate tournament held at a midwestern university in January, 1989, an inexperienced and experienced judge were present in each round of debate. In the letter of invitation, it was announced that two judges would be present in each round, and teams would be advanced on the basis of 12 ballots. Therefore, debaters and coaches knew prior to the tournament that an inexperienced person would be judging alongside of an experienced person.

For the purpose of this study an experienced judge was a person actively involved in coaching debate, who had competed in the event, who had past coaching experience, or who had taught a course in the field. Ability to flow was not a criterion used to distinguish categories. An inexperienced judge was defined as one who did not have the above qualifications, and these judges ranged from professors not involved in competitive debate to local

business owners to housewives. Each judge was required to fill out a judging philosophy form, which was then photocopied and distributed to all the debaters.

There were 28 debate teams and 168 ballots from a pool of 62 judges, 39 inexperienced and 23 experienced. One ballot was randomly selected from each judge in the preliminary rounds of the Open Division of the CEDA debates. Standard CEDA ballots were used in every round.

To answer Question #1, the number of times the experienced and inexperienced judges agreed or disagreed on their decisions was tabulated. Outrounds of spring and fall CEDA tournaments across the country during the 1989-1990 season provided one standard of comparison regarding how often judges typically agree. Another comparison was made with a midwestern state high school debate tournament where two experienced judges were used in each preliminary round.

The answer to Question #2 was found by tabulating speaker ratings from the standard CEDA ballots (30 point scale).

In order to answer Question #3, a content analysis of all the preliminary round ballots was conducted. Content analysis was selected because it allowed the authors to accurately conceptualize the round as well as clearly establish the boundaries and goals for the research (Krippendorff, 1980). Thus, content analysis of the comments on a ballot assisted in determining which "elements, assertions, or arguments" influenced the decisions of judges so that an inference regarding differences between inexperienced and experienced judges would be possible (Krippendorff, 1980, p. 27). This process made the establishment of a list of themes used in the reasons for decisions/comments section of the ballot possible.

According to Holsti (1969), in order to establish the validity of a content analysis, the following steps should be taken. The themes must "adequately reflect the investigator's research question" (p. 95). In order to compare the reasons for decisions by experienced and inexperienced judges, predetermined themes were not superimposed but constructed from the total sample of 168 ballots.

Another step is that "all relevant items in the sample of documents under study must be capable of being placed into a category" (Holsti, 1969, p. 99). By having each comment on the ballot represented by a theme, the element of exhaustiveness was met.

Themes must also be "mutually exclusive...no content datum can be placed in more than a single cell" (Holsti, 1969, p. 99). Every ballot was examined to insure that all possible themes were detected and each theme was discussed until its parameters were agreed upon to safeguard that one comment could only fit into one theme (Weber, 1985; Holsti, 1969).

The next major phase of the content analysis involved an independent examination of the 62 randomly selected ballots, one from each judge. Since critics judged a different number of rounds, ballots were randomly selected so that each ballot "has the same chance of being represented in the collection of sampling

units [which]... assures that there is no bias in the inclusion of units in a sample" (Krippendorff, 1980, p. 66).

Each author analyzed the comments on the ballots and identified the themes to which they belonged. To generate a final composite of each judge, independently established lists were compared and disagreements were resolved by negotiation, which sometimes included re-examining the ballot, until consensus was reached. This provided an additional assurance that each comment was assigned to only one theme.

A content analysis should also insure that "conceptually different levels of analysis must be kept separate" (Holsti, 1969, p. 100). The present analysis was based upon one level because it only involved determining to which themes judges' comments belonged. The number of times a particular theme was used was not tabulated; instead, a theme was noted regardless of how many times it was present on the ballot.

Now that the validity of the methodology has been established, the second vital element is reliability. To secure reliability, defined by Krippendorff (1980) as "the degree to which a process can be recreated under varying circumstances at different locations using different coders," the authors independently coded each randomly selected ballot and then compared their results. This consistency is called "reproducibility... An example is when two or more individuals apply the same recording instructions independently on the same set of data" (p. 131).

Finally, the findings were transferred to computerized forms. Each theme was treated as a variable. A one (1=yes) or a five (5=no) was marked to indicate whether or not the judge relied upon that theme. Comments by inexperienced and experienced judges were compared by using Chi Square.

RESULTS

Question #1

Sixty-four percent of the time there was agreement between the inexperienced and experienced judges in the rounds, and 36% disagreement.

Table 1

CHART OF JUDGES DECISIONS

	AGREEMENT	DISAGREEMENT
Rd. 1	20	8
Rd. 2	14	14
Rd. 3	12	16
Rd. 4	26	2
Rd. 5	16	12
Rd. 6	20	8
TOTALS	108	60
	64% Agreement	36% Disagreement

Question #2

The tabulation of speaker ratings revealed findings similar to Miller (1969); inexperienced judges ($\underline{M}$ = 23.53) assigned higher speaker points than did experienced judges ($\underline{M}$ = 22.59), $\underline{t}(248)$ = 3.76, $\underline{p}<.01$. (The $\underline{N}$ of 248 is accounted for by the fact that each of the 62 judges provided 4 ratings.)

Question #3

Rather than superimpose any categorization for the comments by either inexperienced or experienced judges, the two coders independently examined all ballots noting every particular reason for decision by each judge. The number of times each label was used was not tabulated, but an exhaustive list of all possible labels was sought. The agreement rate between the two coders concerning "labels" for the reasons for decisions present on all the ballots was 88%. On occasion different labels were used for similar themes. For instance, one coder used "judging paradigm" while the other used the designation of "judging philosophy." Negotiation was used until consensus was reached resulting in the generation of 17 themes. Following a second reading of all the ballots and further discussion, the authors agreed that two themes needed to be combined. "Jargon" was placed into the "General and Miscellaneous" theme and "Reasoning" into "Strength and Weakness of Argument."

Table 2

Themes From Ballots

Delivery	Topicality
Evidence	Organization
Judging Philosophy	Specific Speaker Delineation
Adaptation	
Personal Views	Specific Case Points
Cross Examination	Strength and Weakness of Arguments
Refutation/Rebuttal	
Offcase/Value Objections	Criteria/Value
General/Miscellaneous	

To further explain the parameters of the themes, definitions and actual comments from the ballots are provided.

Table 3

Delivery	rate, rudeness, language choice, comprehension, & nonverbals	"finally-a debater who speaks at a normal rate of speed...good job of reading evidence so I could understand it"
Evidence	source citation, recency of evidence, & credibility of source.	"negative evidence not specific to waiting period" and "much evidence presented"
Judging Philosophy	mentioning philosophy itself	"I am offended that the negative has chosen to ignore my judging philosophy"
Adaptation	comments referring to adapting to both judges	"Adapt to your entire audience" and "given the fact we have a judge in this room that is judging his first round"
Cross Examination	any reference to cross examination	"Strong 2CX-that gave a slight advantage to the negative team"
Topicality	any reference to topicality	"I believe that your case is topical"
Organization	jump around too much, order, arrangement, structure	"your organization was very weak" & "each team could do a better job of previewing and signposting for us"
Specific Speaker Delineation	not just reference to aff. or neg. side but 1AC, 1NC, etc. or speaker's name	"2 aff-good point-clarifying purpose of waiting period" & "(name)-you don't need to read all those blocks, ugh"
Specific Case Points	reference to outlining, enumeration of contentions	"Cont I & II fall" & "Obs. 2 is labeled 'Effective Handgun Limitations Are Impossible'"

Strength and Weakness of Arguments	reasoning, analysis, evaluation of arguments	"neg missed the Big Boat- They never linked VO's to aff case" & "you need to understand the difference b/t causality & correlation"
Refutation/ Rebuttal	responsive to, inadequate answer, clash, drops, countered	"be careful not to drop Observations in 1NC" & "effective clash on C-E relationship of h. guns & death w/ wise neg arguments for counter causes"
Off-Case/VO's	reference to by name or disads	"OFF CASE IN THE END MADE NO DIFFERENCE" & "VO #3 weapons shift flips case"
Personal Views	indication of a personal preference regarding a particular claim	"While negative made the better case, in terms of style & content, positive was where my heart is!" & "I voted neg. because aff. is banning hand guns...there are decent people who enjoy hunting, collection etc."
Criteria/Value	reference to by name	"gave standards & criteria for judging" & "the value argument was never central to the round"
Misc/General	any comment that did not fit the previous themes	"judges are not morons" & "I'd dress up & look the part if I were you" & "Will you stop snapping your fingers!!!" & "This is good tight round"

To further answer Q3, one randomly selected ballot from every judge was analyzed and comments were coded using the 17 themes. A comparison of the independent coding resulted in 80% agreement and a Scott's Pi of 0.78. Again, discussion was used to resolve differences. Thus, each judge was represented by a composite of the themes he or she used on the ballot. The composite was placed on scantron sheets so that a Chi Square analysis could be performed. Table 4 provides a summary of the results of the statistical analysis regarding the relation between a judge's experience and the use of a particular theme.

In addition to comparing how each theme was used by judges, an overall comparison was conducted. Each theme was treated as one unit, and a t-Test was performed to determine if a statistical difference existed regarding reasons for decisions. A score was calculated for every judge depending upon the number of themes he or she used. For instance, for each theme used, a score of "1" was assigned; whereas, if a theme was not used, a score of "5" was given. Thus, the higher the mean, the fewer themes used. There was a significant difference between the 39 inexperienced judges ($\underline{M}$ = 26.48) and the 23 experienced judges ($\underline{M}$ = 23.17), $\underline{t}$ = 5.45, adjusted df= 49.8, p<.00.

Table 4

Name of Theme	Chi Square	Probability	Use of Theme Experienced		Use of Theme Inexperienced	
			Yes	No	Yes	No
Delivery	X2 (1, N=62)=.16	p<.69	20	19	13	10
Evidence	X2 (1, N=62)=3.64	p<.06	14	25	14	9
Judging Philosophy	X2 (1, N=62)=.60	p<.44	1	38	0	23
Adaptation	X2 (1, N=62)=3.99	p<.05	2	37	5	18
Cross Examination	X2 (1, N=62)=7.65	p<.01	6	33	11	12
Topicality	X2 (1, N=62)=5.58	p<.02	4	35	8	15
Organization	X2 (1, N=62)=1.0	p<.32	9	30	8	15
Specific Speaker Delineation	X 2 (1 , N=62)=11.46	p<.00	19	20	21	2
Specific Case Points	X 2 (1 , N=62)=11.26	p<.00	0	39	6	17
Strength/ Weakness of Argument	X 2 (1 , N=62)=13.03	p<.00	20	19	22	1
Refutation/ Rebuttal	X2 (1, N=62)=6.29	p<.01	16	23	17	6
Off Case/VO's	X2 (1, N=62)=7.99	p<.01	1	38	6	17

Personal Views	X2 (1, N=62)=.02	p<.89	2	37	1	22
Criteria/ Value	X2 (1, N=62)=4.21	p<.04	5	34	8	15
General/ Misc.	X2 (1, N=62)=4.54	p<.03	18	21	17	6

DISCUSSION

Contrary to what has been suggested by several authors, decisions by inexperienced judges do not differ radically from those of experienced judges. This finding becomes impressive when one compares it to the only available data regarding agreement among collegiate judges and the decisions rendered in elimination rounds. Data distributed by CEDA regarding the out rounds of 117 tournaments in 1989-90 throughout the United States, reveals that out of 1,598 decisions, only 829 or 52% were 3-0 decisions. Thus, 48% of the time there is a 66% agreement rate regarding judges decisions where typically experienced judges are involved. In addition, 78 ballots from preliminary rounds at a 1990 midwestern state high school tournament provide further data for comparison. The results showed that when two experienced judges were present in a round there was agreement only 72% of the time and disagreement 28% of the time. The agreement rate between inexperienced and experienced judges when compared to that of two experienced judges reveals an 8% difference. This difference can be accounted for considering that expert judges were involved. Thus, the agreement/disagreement rate between inexperienced and experienced judges appears to be close to the norm.

Since there is no absolute or "right" agreement percentage rate between or among judges, even when they are experts, the small amount of disagreement that was found should not be used as justification to dismiss the decision of the inexperienced judge. When hearing students complain that they "lost because it was a lay judge," coaches can confidently explain that this complaint is rarely valid.

The statistically significant difference that was found in Question 2 suggests that experienced judges do not assign higher speaker points due to expectations regarding a particular debate team, as is claimed by Phillips (1985). He argues that speaker points below 24 are rarely given. Yet, the **average** speaker points given by experienced judges was 23.17. Instead, inexperienced judges assign higher speaker points and thus might be less critical and more easily impressed.

Concerning reasons for decisions offered by judges, numerous authors (Durkee, 1985; Marrs, 1990; and McGough, 1988) claim that using inexperienced judges is a means for controlling "spit and spread" tactics and increasing the "real world" communication training. Conversely others (Phillips, 1985; Flood & Cripe, 1982; and Gotcher, 1985) note how inexperienced judges rely far too

heavily on the communication aspect of the round. The results of this study indicate, however, that neither of these claims is true. Inexperienced judges do not comment significantly more on delivery than experienced judges. Participants in this debate tournament were warned that an inexperienced judge would be in each round and thus knew a need to adapt existed. Also, the absence of any comment regarding delivery does not establish that it was not influential at all in the judges decision. But, based upon all the available data to students and coaches, there was no difference. Nor was a significant difference found regarding evidence, judging philosophy, or organization. One indictment of inexperienced judges is that they are more apt to accept weakly supported claims. The results deny this charge and indicate that inexperienced judges stress the amount and quality of evidence as do experienced judges. One might expect that experienced judges, who have filled out numerous judging philosophies, would refer to their philosophy more than inexperienced judges who rarely, if ever, have filled one out. However, no significant difference was found. One implication of this will be discussed under adaptation. Finally, although it is commonly believed that inexperienced judges will be more concerned with organization, this difference does not exist. Both categories of judges are equally concerned with the arrangement and structure of the arguments.

A significant relationship was found in the adaptation theme--experienced judges were more likely to make comments about the necessity of adapting, especially to the inexperienced judge in the round. Although experienced judges did not comment significantly more about adapting based upon judging philosophies, they were concerned with debaters accommodating both judges in the round. Also, as noted by Hanson (1989), inexperienced judges often lack the confidence to demand that debaters adapt to them or might not realize the importance of the adaption aspect of a debate round, thus not commenting about it. A significant difference was found in comments made about cross-examination ($p<.01$). Perhaps the experienced judges make more comments because those in the activity pay more attention to what is happening in cross-examination, and the inexperienced appear not to be as influenced by it. Experienced judges might see it as a strategic time to exercise critical thinking skills, and they reward the quality of questions utilized by the debaters. On the other hand, the lay judges may have viewed cross-examination as more for the debater's sake of clarification rather than for the judge.

Topicality is another area in which a significant difference was found ($p<.02$). Experienced judges commented on it over 50 percent of the time, while inexperienced used it approximately 11 percent of the time. Because the word itself is an "in-house" or jargon word, experienced people were more likely to use it.

The significant difference between inexperienced and experienced judges regarding specific speaker delineation ($p<.00$) was not surprising since 21 of 23 experienced critics address each speaker. These judges are more aware of the particular duties and perhaps are in the habit of structuring the ballot speaker by

speaker. In other words, experienced judges are not just seeking to provide justification for the decision, but guidance regarding strategy and specific flaws or strengths in particular speakers. Thus, experienced judges provide a more accurate account of what happened in the round. Inexperienced judges, on the other hand, when directing comments to speakers, do so by name rather than position or speech. They stress personal interaction with the speaker and are more concerned with the speaker's credibility. Whereas, experienced judges focus more upon strategic choices of a particular speaking position.

Similar to specific speaker delineation, there was a significant difference regarding specific case points ($p<.00$). Not a single inexperienced judge whose ballot was randomly selected commented about specific case points by outline label or enumeration. Inexperienced critics look at the totality or the big picture of debate versus the utility of Contention I, subpoints 1 and 5 under A. This warns the forensic community that seasoned judges may be overly concerned with identification tags. Also, since lay judges often did not flow, they did not "just pull tags" but required the debaters to recap the essence of a claim and refute it or expand upon it. Or, the absence of a flow did not allow inexperienced judges to associate specific case points with particular arguments. Instead, inexperienced judges seem to respond more like public audiences in general--to large ideas conveyed through effective warranting. (Cox, Jensen, Wheeler & Fulton, 1991).

One of the most important findings of this research concerns comments regarding strength and weakness of arguments ($p<.00$). When it comes to evaluating argumentation, experienced judges reflect more upon the reasoning and analysis of the debaters when explicating their decision. Although some might argue that the reason for this difference in emphasis is because experienced judges know more about the topic, the tournament at which the data was collected was one of the first in the semester. Earlier it was argued that inexperienced judges look at the overall picture, but this type of analysis falls short in the achievement of the goal of critical thinking. Experienced judges are more analytical in their comments and are less tolerant of undeveloped arguments. If this finding were the only basis to determine whether or not an inexperienced critic should be used, then exclusion of lay judges would be warranted.

Another theme that experienced judges used significantly more than inexperienced was refutation/rebuttal ($p<.01$). Those involved in the activity are concerned with how well the debaters can clash and counter their opponents' arguments. Perhaps it is enactment regarding the burden of rejoinder. On the other hand, seasoned judges may be obsessed with countering every specific point regardless of its merit. For instance, the judge might state, "1NR you dropped subpoint 7, therefore I vote affirmative." Inexperienced judges appear to listen more to the totality of the arguments and worry less about specific points of contention. Taking substantial notes or flowing the debate round might be a

factor that enters into these results. As Phillips (1985) and Freeley (1986) have argued, a judge should flow the debate round. If inexperienced judges do not flow, or flow well, then they would not be able to remember what exact points were and were not refuted. Some of the educational value of the comments provided by an inexperienced judge is diminished if the comments are not complete in nature.

Only 1 out of 39 inexperienced judges commented on the theme off-case/VO's ($p<.01$). As mentioned previously, flowing might be a factor. This particular theme also is related to identifying the format of debate; typically second negative deals with off-case or value objections. Holm (1940) and Balthrop (1983) insist that a judge must be familiar with debate procedures and participant obligations. An inexperienced judge might not understand what a "VO" is and thus not feel comfortable commenting on it. Again, the educational dimension is reduced if comments are not thorough because a specific term is not understood. Conversely, if these words are jargon terms that only those involved in forensics understand, debate is then a closed system as McGough (1988) and Decker and Morello (1984) posit, and an inexperienced judge helps bridge the artificial environment to the "real world" (Durkee, 1985). While most technical activities are bound by field dependent language, one of the founding principles of CEDA is to provide debate appropriate for public consumption (Jensen & Preston, 1991; The Vision and the Reality, 1991; Hollihan, Riley & Austin, 1983).

Contrary to what Srader argued that lay judges intervene and weigh the debate round against their own experiences, no statistical difference was found between the groups of judges regarding comments containing personal views. Perhaps one would expect inexperienced judges to express their personal beliefs about the topic more often, but the data does not confirm this expectation. This finding might help account for the fact that the decisions at the end did not differ a great deal.

Similar to off-case/VO's, experienced judges commented more on criteria/value ($p<.04$) than did inexperienced judges. Since these terms are predominantly associated with CEDA debate, these results are not at all surprising. Anyone outside the debate realm would be unfamiliar with these terms and not use them on the ballot.

Finally, a significant difference was found in the category of miscellaneous comments ($p<.03$). Experienced judges write more general comments because they feel free to do so and have a broader knowledge base from which to draw. They might know the debaters personally or have judged them numerous times. However, inexperienced judges might not think that the miscellaneous/general comments are appropriate, and perhaps these judges felt uneasy about making comments outside the realm of the debate round.

In an overall comparison, experienced judges used significantly more themes than did the inexperienced ($p<.00$). Since these themes encompass the essence of what is necessary to accomplish the goals of debate--critical thinking, communication skills and academic advancement--experienced judges ought to be used predominantly over

inexperienced judges. Those within the activity better serve the educational functions due to their breadth of comments.

If only delivery, organization and evidence are the essential elements to achieving the ends of debate participation, then inexperienced judges should be used just as often as experienced. However, no one claims that these three factors alone are sufficient to insure that debate accomplishes its objectives. As long as the primary function of the judge is to render a decision, then inexperienced judges can be used with confidence. The educational dimension of debate will then need to be supplemented after the tournament by the coach. However, if a judge is to function as both an adjudicator and an educator, then assignment of inexperienced judges should be limited. Implications for Future Research

Additional generalizations regarding the issue of using inexperienced and experienced judges in debate tournaments need to be based upon verifiable data rather than anecdotal observations. This study illustrates that anecdotal information is often misleading. For instance, the common belief that inexperienced judges focus far more on the element of delivery was not verified by the data. Thus, scholars need to be cautious not to make hasty generalizations.

Future research could expand upon this study, which has determined that a difference exists between categories of judges, and include counting the number of times a judge relied upon a theme. Such an extension might provide insight into the subtleties of the differences and/or similarities in reasons for decisions. Expansion could also take place by increasing the sample size and using ballots from more than one tournament. Another change might include not telling debaters of the presence of inexperienced judges. Since debaters were primed to adapt, information gathered might have been affected. Additionally, a contingency analysis between themes used, side voted for, and win/loss decisions might be attempted.

Finally, a comparison of inexperienced judges who receive some training prior to judging to those who do not, could provide needed insights regarding what should be included in judge training. In fact, the value of training itself might be assessed.

Because of the scarcity of information upon which conclusions can be drawn, hopefully future research will take into consideration the use of inexperienced and experienced judges.

REFERENCES

Balthrop, V. William. "The Debate Judge as 'Critic of Argument': Toward a Transcendent Perspective." Journal of the American Forensic Association 20 (1983): 1-15.

Brydon, Steven R. "Judging CEDA Debate: A Systems Perspective." CEDA Yearbook. Ed. Don Brownlee. Cross Examination Debate Association, 1984: 85-88.

Burgoon, Judee K. "Evaluation Criteria as Predictors of Debate Success." Journal of the American Forensic Association 12 (1975): 1-4.

Colbert, Kent R. "A Quantitative Analysis of CEDA Speaking Rates." Speech Communication Association Convention. New Orleans 3 Nov. 1988.

Colbert, Kent R. "Criticizing Debate Critics." Speech Communication Association Convention. San Francisco, 19 Nov. 1989.

Cox, E. Sam, Scott L. Jensen, Chistopher Wheeler, and Michael Fulton. "An Objective Analysis of One Audience Debate." Paper presented at the 7th SCA/AFA Conference on Argumentation, Alta, UT, 1-3 Aug. 1991.

Cox, J. Robert. "A Study of Judging Philosophies of the Participants of the National Debate Tournament." Journal of the American Forensic Association 11 (1974): 61-71.

Davis, Bill. "In Defense of the Lay Judge." Rostrum Feb 1990: 17-18.

Decker, Warren D., and John T. Morello. "Some Educational Difficulties Associated with Mutual Preference Debate Judging Systems." Journal of the American Forensic Association 20 (1984): 154-61,

Durkee, John. "Judging: A Response." National Forensic League Conference. Kansas City, MO 10 Aug. 1985.

Flood, Royce E., and Nicholas M. Cripe. Modern Competitive Debate. Amherst, NY: Poppenberg's Press, 1982.

Freeley, Austin J. Argumentation and Debate Critical Thinking for Reasoned Decision Making. Belmont, CA: Wadsworth, 1986.

Friedman, Robert P. "Reflections of an Incompetent Judge." Journal of the American Forensic Association 8 (1971): 123-26.

Ganer, Patricia M. "The 'Emperor Phenomenon': The Necessity of Critic Responsibility." Argument and Critical Practices: Proceedings of the Fifth SCA/AFA Conference on Argumentation. Ed. Joseph W. Wenzel. Annadale, VA: Speech Communication Association Convention, 1987. 387-93.

Glass, David. "Debate Should be Specialized." Rostrum Mar. 1990: 19.

Gotcher, James M. "Who Should Judge CEDA Debate?" Southern Speech Communication Association Convention. Winston-Salem, 11 Apr. 1985.

Hanson, Colan T. "Strengthening the Tournament Experience: Developing A More Explicit Set of Expectations for our Judges." Developmental Conference. Pi Kappa Delta National Convention. St. Louis, 22 Mar. 1989.

Henderson, Bill, and David L. Boman. "A Study to Determine if Debate Judges' Judging Philosophy Statements are Consistent with their Later Related Ballot Statements." Journal of the American Forensic Association 19 (1983): 191-98.

Herbeck, Dale A. "The Nature and Scope of Evaluative Criteria: An Argument for Nonintervention." Central States Speech Association Convention. Chicago, 12 Apr. 1984.

Hill, Sidney R., Jr. "Prestige as a Factor Influencing the Validity of Decisions in Intercollegiate Debate." Southern Speech Communication Journal 40 (1975): 351-64.

Hollihan, Thomas A., Patricia Riley, and Curtis C. Austin. "A Content Analysis of Selected CEDA and NDT Judges' Ballots." Argument in Transition: Proceedings of the Third Summer Conference on Argumentation. Eds. David Zarefsky, Malcolm O. Sillars, and Jack Rhodes. Annandale, VA: Speech Communication Association, 1983. 871-82.

Holm, James N. "A Philosophy of Judging Debate." Southern Speech Communication Journal 1 (1940): 5-11.

Holsti, Ole R. Content Analysis for the Social Sciences and Humanities. Reading, MA: Addison-Wesley, 1969.

Jensen, Scott L. & C. Thomas Preston. "Intent and Practice: An Examination of Mission and Practice in Cross Examination Debate Association Debate." Speech and Theatre Association of Missouri Journal 1991 (forthcoming).

Klopf, Donald W. Coaching and Directing Forensics. Skokie, IL: National Textbook, 1982.

Krippendorff, Klaus. Content Analysis: An Introduction to Its Methodology. The Sage COMMTEXT Series 5. Beverly Hills: Sage, 1980.

Lybbert, Blair. "What Should be the Goals of High School Debate?" National Forensic League Conference. Kansas City, MO 8 Aug. 1985.

Markgraf, Bruce. "Debate Judging and Debater Adaptation: A Reply." Journal of the American Forensic Association 3 (1966): 37-39.

Marrs Dan. "Lay Judges Revisited: The Perspective from the Other Side of the Ballot." Rostrum Apr. 1990: 16.

Mayer, Michael E. "Changing the Debate Format: A Mechanism to Control the Spread." Western Speech Communication Association. Fresno, CA 16 Feb. 1985.

McCroskey, James C., and Leon R. Camp. "Judging Criteria and Judges's Bias in Debate." Journal of the American Forensic Association 3 (1966): 59-62.

McGough, Michael. "Pull it Across Your Flow." New Republic 10 Oct. 1988: 17-19.

Miller, Arthur B. "Instruction of Debate Judges: A Case History." Journal of the American Forensic Association 6 (1969): 24-26.

Nicolai, Michael T. "The Use and Abuse of the Hired Judge Pool." Speech Communication Association Convention. Denver, 10 Nov. 1985.

Patterson, J. W., and David Zarefsky. Contemporary Debate. Boston: Houghton Mifflin, 1983.

Pearce, W. Barnett. "Communicating the Reasons for Decision by the Ballot." Journal of the American Forensic Association 6 (1969): 73-77.

Phillips, Leslie. "The Debate Judge as Educator." National Forensic League Conference. Kansas City, MO 10 Aug. 1985.

Richardson, G. David. "Response to: 'The Debate Judge as Educator' by Leslie Phillips." National Forensic League Conference. Kansas City, MO 10 Aug. 1985.

Rowland, Robert C. "The Debate Judge as Debate Judge: A Functional Paradigm for Evaluating Debates." Journal of the American Forensic Association 20 (1984): 183-93.

Rowland, Robert C., and Scott Deatherage. "The Crisis in Policy Debate." Journal of the American Forensic Association 24 (1988): 46-50.

Sanders, Gerald H. Introduction to Contemporary Academic Debate. Prospect Heights, IL: Waveland Press, 1983.

Schunk, John F. "A Quantitative Study of Decision-Making in Academic Debate." Journal of the American Forensic Association 7 (1971): 271-77.

Scott, William A. "Reliability of Content Analysis: The Case of Nominal Scale Coding." Public Opinion Quarterly 19 (1955): 321-25.

Sodikow, Richard B. "Pogo the 'Possum Lives, or We Have Met the Enemy, and They Still Are Us." National Forensic League Conference. Kansas, City MO 8 Aug. 1985.

Srader, Doyle. "The Real Value of Judge Adaptation: A Reply to Bill Davis and other East Kansas Debate Coaches." Rostrum May 1990: 9-10.

The Vision and the Reality: A Conversation with CEDA Presidents. Videocassette. Prod. Western Washington University, February 1991. 120 min.

Ulrich, Walter. Guidelines for the Debate Judge. Kansas City, MO: National Federation of State High School Associations, 1985.

Verderber, Rudolph. "Judges' Criteria and Debater Adaptation: Empirical Evidence." Journal of the American Forensic Association 5 (1968): 28-30.

von Moltke, Henry. "Decision Debating: A Judge's Point of View." Journal of the American Forensic Association 1 (1964): 98-100.

Wallmark, Michael. "Back to Basics in Debate." National Forensic League Conference. Kansas City, MO 8 Aug. 1985.

Weber, Robert Philip. Basic Content Analysis. Quantitative Applications in the Social Sciences 49. Beverly Hills: Sage, 1985.

MEANING AS LANGUAGE USE: THE CASE OF THE LANGUAGE-LINKED VALUE OBJECTION[1]

Ken Bahm
Gonzaga University

"When I use a word," Humpty Dumpty said in a rather scornful tone, "it means just what I choose it to mean - neither more nor less."
"The question is," said Alice, "Whether you can make words mean so many different things."
-Lewis Carroll, Through the Looking Glass

In Wittgensteinian terms, Humpty Dumpty would not have been considered much of a linguist. By presenting meaning as a simple choice made and imposed by the language-user, this view seriously endangers any possible concept of language as an active and consensual process.

While it is perhaps unsurprising to recognize philosophical naivety in the characters of Carroll's book, it should be more surprising to recognize Humpty Dumpty's analysis in the defining norms of contemporary academic debate. When we tell an affirmative team that they have a "right to define," we are telling them that they, as language users, can select meaning -- can choose what a word will mean, neither more nor less. The legitimate question to be asked -- Alice's question -- is whether words can, in fact, be made to mean at all.

This paper focuses on the debate community's conception of meaning. I suggest that seeing meaning as something that an affirmative team, or any team, has an open right to pre-select is dramatically inconsistent with modern philosophies of language, specifically with Wittgenstein's conception of meaning as use. The application of a more functional perspective on meaning will present academic debate with a much less problematic method of determining meaning.

Certainly, it is justified for the philosophy of language to intrude on the world of academic debate. Debate is an activity intimately involved in communication, not an activity separate from other communication disciplines. Goodnight (1981) and Hingstman (1983) have argued that the practice of forensics should forge a closer relationship to communication theory. A broader and more informed perspective can be gained by promoting more relationships between forensics and other communication-related theories.

Specifically, modern theories of meaning should be applied to academic debate. Since debate is most fundamentally a thinking and communicating activity, concepts of meaning which are

[1].An earlier draft of this paper was presented at the 1990 Speech Communication Association conference in Chicago Illinois. The author would like to thank several debaters whose advocacy of this position has furthered his own understanding of it: David Devereaux, Jennifer Dunn, Marty Horn, Joe Sullivan, and Mark West.