This essay critiques the theoretical perspectives, research design and analysis, and interpretation and implication of Keogh, Barnes, Joiner, and Littleton's paper "Gender, Pair Composition and Computer versus Paper Presentations of an English Language Task."

Theoretical Perspective

The authors' conceptual framework is well laid out in some theoretical areas, but not in others. For instance, the authors clearly and logically explain that there is an increase use of computers in schools to support collaborative learning; there are gender inequalities in favor of males in computer attitudes, use, knowledge, and experience; and these inequalities may influence the performance in favor of males over females. On the other hand, the authors do not logically argue the role of the type of collaborative learning tasks when explaining the mixed findings for the differential performance of same- and mixed-gender groups in relation to the current study. As shown in the analysis, the authors' do not manipulate or consider this factor for their research design, instead the task remained the same, but the medium—computer or paper—was manipulated. The framework should instead include how performance is different for different genders across mediums. Furthermore, it would have been even better if the authors included how performance was different across mediums for mixed- and same-gender groups. Lastly, the framework does not well address the link between dominance and perceived expertise. That is, it is argued that in an interaction the one who is perceived as more expert is more dominant and the one who is perceived as less expert is less dominant. It is unclear how the authors ruled out other explanations of dominance in favor of perceived expertise. For example, it may be possible that boys are genetically more domineering in

interactions and girls more passive in interactions. Other explanations such as genetics, socioeconomic, or prior experience should be ruled out in their framework.

Nevertheless, if the authors were able to clarify these latter points, it can be argued that this study is fairly important. That is, it would be interesting to understand the consequences of schools moving towards using computers in teaching more frequently, especially in the area of collaborative learning, because as the authors argue, different genders have different attitudes and experiences with computers. If this is true, then how do these differences affect the interactions (and ultimately their performance) of the often-used mixed- and same-gender groups? Ultimately, if the study addresses this question, educational policy and practices would then be affected because it may no longer be fair to place a student in a same- or mixed-gender group knowing that a certain genders behave differently (which future research can investigate if this causes a different group to have a higher level of performance).

In addition, the authors' do provide a moderately effective link between their points and relevant theory and prior research. For instance, several studies are cited for evidence of gender inequalities in computer attitude, use, and expertise. Furthermore, although I mentioned that I feel the authors should provide more research on the performance differences between computer and paper mediums for mixed- and same-gender groups, the authors provide an excellent overview of the mixed findings on performance differences for different gender groups on different task types. In addition, despite my feelings that there should be additional arguments as well as research behind the reasoning for concluding dominance in interactions is explained solely by perceived

expertise, the authors nevertheless provide a fairly adequate presentation of previous research on perceived expertise.

Lastly, the authors provide a fair purpose statement and hypothesis. However, it would have been more informative if they specifically stated their research questions so that the reader would better understand the motivation behind their analyses. For example, the authors conduct several analyses of variance on verbal interaction. It would have been more meaningful, if they would have stated "Do different genders verbally interact differently across computer and paper mediums for same- and mixed-gender groups?"

Research Design and Analysis

The authors use a three-way mixed factor analysis of variance with two between subjects and one within subject factor. This is appropriate for their needs in comparing means across the various combinations of grouping variables. However, the authors run several analyses with possibly correlated dependent variables (e.g., related forms of utterances and related types of manipulations). To control for the correlations and inflation of Type I error rate when conducting multiple tests, it may be more appropriate to conduct a multivariate analysis of variance. In addition, although no significant results were found, the authors should probably also have used a multivariate analysis of variance to control for the inflation of Type I error rates when conducting the multiple independent t-tests for order effects. For follow-up contrasts, they could use a Bonferroni or Tukey correction.

Furthermore, the authors' claim that the West-London based school was socially mixed. However, it is not clear how the school was chosen; hence, if the school was not

randomly selected, then there is low generalizability beyond samples similar to those found in West London with similar characteristics such as the overall socioeconomic status and ethnic diversity. The study would be more complete if details on how the school was sampled were revealed.

Although the authors implement a mixed order presentation to remove any order effects, the study's procedures and materials are fair (at best). For instance, it is not clear if random assignment was conducted. It seems likely that the assignments were not random because the authors mention that they placed the children into groups with their consultation. If this suggests no random assignment, then the internal validity of the study may be hindered. It also seems likely that the directions used in the experiment may have hindered the internal validity. That is, it seems that the experiment administrators discussed the purpose of the experiment with the children. Consequently, findings may be biased because some children could have sabotaged the scores if they knew the purpose of the observations. Additionally, it is unclear if one or multiple student pairs were observed simultaneously by one or multiple raters. If one rater observed multiple pairs, then some coding is likely lost, hindering the reliability of the scores. On that same note, although coding explanations were provided for the types of utterances, the authors did not provide similar explanations for the forms of manipulations that were coded. For completeness, this should have been provided.

In terms of the appropriateness and quality of the measures, the authors provide incomplete evidence that measures were reliable. That is, the authors inform the reader that an independent coder analyzed 25% of the transcripts and agreed with all but three utterances, but the rationale of choosing those 25% and only 25% is unclear. Although

the observed reliability may have been high, a larger random selection of transcripts could have given a more precise estimate. Furthermore, it is difficult to argue the independence of the raters because no information on who and how the coder was chosen was provided. Lastly, other forms of rater agreement measures could have been used to provide a better picture of the reliability of the scores. For example, kappa's coefficient, agreement indices, and correlations could have been used.

As for the adequacy of the data analyses, in addition to the possibility of using a multivariate analysis of variance, the one major confusing statistical interpretation is that the authors interpreted the lower order interactions—after finding a higher order interaction. Because of the higher order interactions, the interpretations of the lower order interaction tests are not accurate. In addition—although not as major of a misinterpretation than the aforementioned, the authors should have probably chosen one alpha-level—before tests were conducted—and determined if their tests are significant or not based on this one alpha-level. Phrases such as more significant or less significant do not make sense—they are either significant or not. If they wanted to discuss magnitude of the differences, then they should have calculated effect sizes.

Lastly, the dependent variables should be checked for violations on the assumption of normality. The dependent variables are counts suggesting a Poisson distribution. Hence, a simple check for a violation of this assumption (as well as other assumptions) should have been conducted before the tests, if none were already conducted.

Interpretation and Implication of Results

The discussion of the methodological and/or conceptual limitations is non-existent. The authors should have discussed this for a well-rounded argument. For example, they could have included the sample limitations, lack of generalizability, and the possibility that not all explanatory variables such as socio-economic status and previous experience were included.

However, the authors do provide consistent and comprehensive conclusions with the reported results. They interpret each finding and relate it back to their original purpose. Unfortunately, as I mentioned earlier, much of the interpretations are incorrect because of the significant higher order interactions.

Nevertheless, the authors relate the results to the study's theoretical base fairly well. They make connections between their results and how it relates to their framework concerning the types of tasks, gender, same- and mixed-gender groups, and perceived expertise. In addition, the authors do an excellent job of discussing how their findings may affect practices in collaborative learning groups. However, as I mentioned earlier, it may be more correct to relate their results to the research on the type of medium the tasks are presented in, and more complete if the argument is provided that dominance in interactions can be solely attributed to perceived expertise.

In my view, the significance, implications for theory, future research, and practice of this study are such: As mentioned above, I feel the significance of the purpose of this study is important because educational researchers and practitioners need to understand the consequences of using computers more frequently in teaching. If there are inequalities to using computers for different genders, then the implications are that changes should be applied to practice so that one gender is not biased against. Furthermore, future research

on computer use and collaborative learning can also focus upon teaching methods or actions that create more equality when using computers.

In summation, this essay argues that Keogh, Barnes, Joiner, and Littleton's paper "Gender, Pair Composition and Computer Versus Paper Presentations of an English Language Task" achieved the following: (1) the theoretical perspectives is generally well laid out except for incomplete links between task medium and perceived expertise, (2) the research design and analysis would be more accurate if inflation for Type I error rates, relatedness of the dependent variables are considered, and lower order interactions are not interpreted if there are significant higher order interactions, and (3) the interpretation and implication of the results are fairly thoughtful by relating their findings to previous research, but the discussion would have been more complete if the limitations were considered.