RESEARCH AND DEVELOPMENT IN ICT IN EDUCATION (ICT.02C2)

Critical review of the article: Warschauer, M. (1996). Comparing Face-to-face and electronic discussion in the second language classroom. CALICO journal 13(2), 7-26

Maria Fountana
Institute of Education
MA in ICT

Word Count: 4,550

AUTUMN TERM 2000

CONTENTS

1. INTRODUCTION

This article, written by Warschauer, attempts to show the differences in terms of equality of participation and of the quality of language used between face-to-face and electronic discussion. The author's theoretical position, methodological procedure and research findings will be discussed within this critical review. The review is comprised of three sections, specialising, localising and generalising, according to the methodological procedure introduced by Brown and Dowling (1998) in their book: "Doing Research/ Reading Research".

2. SPECIALISING

2.1) THE PROBLEM

The title of Warschauer's article provides an indication of the general theoretical and empirical area around which his research will revolve. Computer assisted language learning will be compared to the traditional face-to-face method of teaching and learning foreign languages. Does electronic discussion foster greater equality of participation than face-to-face discussion? This is the main question that constitutes the "problem" in Warschauer's study and initiates his research. More precisely, the author is interested in scrutinising the equalising role of a computer conference system as a medium for discussion, and also in comparing face-to-face and electronic discussion in terms of a) equality of participation in discussion (level of participation) and of b) the quality of the language used (form of participation) in both environments, within an ESL classroom. The equality of participation and quality of the language use can be identified as concept variables in Warschauer's study. He is also interested, at least in an early stage, in relating his findings to various attributes, such as gender, nationality, age, language ability, student attitude and time in the U.S.

The problem in Warschauer's article is presented in detail through five research questions. He is interested in examining: a) what are the chances for students' equal

participation in discussions held both electronically and traditionally, b) who benefits more from equal participation in terms of gender, nationality, age and language proficiency, c) what are the students' attitudes towards participating both in electronic and face-to-face discussion, d) how is language lexically and syntactically influenced in both environments and e) what other differences occur in the language use and interaction style in the two modes. These research questions enable the author to operationalise his problem, but further discussion on operationalisation will be conducted at a later stage of this critical review.

What should be mentioned at this point is that, from a very early stage in his study, the author adopts a theoretical position, a hypothesis. According to his hypothesis, computer mediated communication influences student participation in an equalising way. Throughout his article he refers to his hypothesis several times, but to intimate each time, that through research findings his claim is confirmed. For example: "The electronic discussions were compared to the face-to-face discussion on two measures of complexity, one lexical…and one syntactic. On both measures, the electronic discussion involved significantly more complex language than face-to-face discussion" (Warschauer, 1996). At the end of his article, the hypothesis is presented as a conclusion: "The findings of this study suggest that electronic discussion may create opportunities for more equal participation in the classroom. Furthermore, this apparently can be achieved without disadvantaging more verbal students"(Warschauer, 1996).

2.2) THEORETICAL FIELD

Sociology and psychology, linguistics and information and communications technology constitute the academic disciplines of the theoretical field of this study. Within the latter, several authors are cited to justify and strengthen the researcher's position. The authorities in this article are not composed only of academics though but of practitioners as well. Historical information is also provided with regard to the development and integration of electronic communication in the teaching of composition and into language teaching. The references to practitioners and to historical information within the article constitute the professional disciplines within

the theoretical field. Brown and Dowling (1998) argue about the theoretical field of educational research, that it "…is not confined to the academic disciplines and certainly not to the traditional academic disciplines. There are also what might appropriately be referred to as professional disciplines".

2.3) THE PROBLEMATIC

The synthesis of the theoretical field is an indication of the disciplinary knowledge within the article. There is a need, though, to narrow down to a more specific field within the general theoretical one, that constitutes the "problematic". Brown and Dowling (1998) define the problematic as "…the denoting of key work, positions and debates within which the research is situated". Warschauer, indeed, cites in support of his position, the findings of relevant experiments. The experiments cited by the author examine the relation between electronic discussion and several variables such as gender (McGuire, T., Kiesler, S., & Siegel, J., 1998, Flores, M., 1990, Selfe, C., 1990- quoted by Warschauer, 1996), social status (Huff, C., & King, R., 1998- quoted by Warschauer, 1996), nationality (Tella, 1992-quoted by Warschauer, 1996), language complexity (Kern, R., 1996- quoted by Warschauer, 1996), members of organisations (Sproull, L., & Kiesler, S., 1991- quoted by Warschauer, 1996).

Additional literature is cited to support the research and to inform the readers about an existing debate in the area. Warschauer cites Hiltz and Turoff (1978/1993), who argue that "computerised conferencing would eventually have dramatical and sociological impacts on various group communication objects and processes". The author shows no special interest in providing more details about this debate, whereas he has been more adequate in presenting the findings from relevant experiments. This could be misleading for the readers[1].

After presenting all the aforementioned literature supporting different parts of his research, the author indicates that, although the findings from many experiments cited

---

[1] A more detailed discussion about the author's readership will be made at a later stage of this critical review.

in this article suggest adequate research has already been made in the field, the latter has comprised only comparisons between whole class discussions, whereas little or no research has been made in relation to small groups "where patterns of interrogation might be different"(Warschauer, 1996). He also indicates that no studies have been conducted to explore and analyse language complexity with ESL students.

3. LOCALISING

3.1) EMPIRICAL FIELD

Warschauer, empirically, intended to explore the comparison of face-to-face and electronic discussion. He chose to focus his research on ESL (English Second Language) students and on small groups within a classroom. His choice is well justified. As one can see from the proposal of several subjects that the author recommends for future study at the end of his article, he has identified several gaps in the field. His study aimed to fill some of these gaps.

3.2) EMPIRICAL SETTING

Brown and Dowling (1998) refer to the empirical setting as "the localized region of the empirical field". Warschauer's empirical setting consisted of sixteen out of twenty students in an advanced ESL composition class, at a community college in Hawaii, who were "randomly" assigned to four numerically equal groups[2]. Four of the students didn't participate in the study since they were absent on that day.

3.3) SAMPLING TECHNIQUES

The quality of a research depends on many parameters. One of these is the sampling technique that a researcher has adopted, and its suitability.

Warschauer has chosen an opportunity sample or otherwise called a convenience (Cohen et al., 2000) sample.

The author was a faculty researcher at the college of Language, Linguistics, and Literature, at the University of Hawaii, from 1994 until 1998. So it is obvious why his

---

[2] The author's decision to "randomly" assign students to groups will be discussed at a later stage of this critical review.

sample can be argued to be an "opportunity sample". It is worth mentioning what Cohen et al. (2000) argue about this kind of sample, which according to them, represents only itself and generalisations should not be made in terms of the wider population. Brown and Dowling (1998) are, though, more flexible in their opinion and argue that "where the empirical setting is defined by an opportunity sample, the validity of generalisation relies on the researcher marking out the continuities and discontinuities between the setting and the empirical field in an ad hoc manner".

The researcher's opportunity sample consisted of students of different nationality (five Filipinos, five Japanese, four Chinese and two Vietnamese), gender (two male and fourteen female students) and age (from nineteen to forty-four years of age). All students had already some experience of the Interchange Computer program (Daedalus) that was used in this study, and had beginning typing skills. The author also provided some more information to his readers, about the criteria[3] that students had to meet in order to be enrolled in this classroom, in the first place.

3.4) RESEARCH DESIGN ISSUES

Adopting an experimental design, the researcher of this study conducted a "controlled experiment" (Warschauer, 1996). Several issues should be discussed about the way in which this design was applied by the researcher.

Firstly, the researcher made his problem operationalisable by expressing it through five research questions[4]. Then, he clarified his intention to observe the effect that the independent variable would have on the dependent variable. The independent variable in Warschauer's study is the medium for discussion and is nominally scaled with two values. These are the face-to-face and the electronic discussion. The dependent variable is both the level and the form of participation[5].

---

[3] All students had received an 11-12th grade in a general English placement test and a passing grade on a writing sample test, and a score of 13.0 in the Nelson Denney reading test or a C grade or better in a beginning-level writing course (English 22).

[4] See page 1.

The author justified his statement that he conducted a "controlled experiment" by claiming that he has been able to "randomly" assign his sample to four groups of four students each.

At this point, this claim needs to be discussed. His main aim was to be able to make a comparison between the groups, when analysing his data, at a later stage. In order to be able to achieve this, it is necessary that the groups have as many identical characteristics as possible. Randomly assigning individuals to groups is a very good way of achieving this. However, there are some limitations. According to Brown and Dowling (1998), there should be "nothing motivating the selection and allocation of a particular person to a particular group. With a numerically large sample, the profile of each group should be similar to that of the population from which the sample is drawn". Is this the case here? As the sample is so small and generally of such a heterogeneous nature, the researcher cannot assert that the groups have been equivalent and therefore the question on whether he has achieved anything by this, needs to be considered. Nevertheless, this fact undermines his claim that he has adopted an experimental design. An alternative strategy that could have been considered by the author is to pair individuals with similar characteristics and to allocate them in different groups, as Brown and Dowling (1998) suggest.

3.5) DATA COLLECTION TECHNIQUES

The researcher used several different techniques to collect his data. He conducted a non-anonymous survey with nineteen general questions, answered on a five-point Likert scale, plus six personal questions to gather information about the students' personal background and attitudes.

There are two points worth raising about this survey in terms of validity. On the one hand, as the author very correctly suggests, the lack of anonymity, although necessary, according to Warschauer, might have influenced the students' answers.

---

[5] Or otherwise mentioned in this critical review as equality of participation and quality of the language use.

Therefore the students might have not answered what they wanted, but what they were expected to answer[6].

On the other hand, this survey needs to be questioned as an instrument used to measure attitude, firstly because the author doesn't define what he means by "attitude". Judith Bell (p.203, 1999) states that every important terms should be defined precisely and Cohen et al. (p.12, 2000) agree with the above statement. "Attitude" is an important term in Warschauer's article, since he considers it to be a variable to be examined. Secondly, the author states that "mean scores and distributions of survey answers were reviewed to determine which questions students had strong opinion on" (Warschauer, 1996). How many of the nineteen general questions in the survey were supposed to examine the students' attitude towards face-to-face and electronic discussion? This is not clear in the article. The author provided the students' answers in only twelve questions. Are these responses mistakenly treated as individual statements? In my opinion, this is the case here.

Data about students' language proficiency were assessed using a Secondary Level English Proficiency Test (SLEP) that tested students' ability in listening and reading comprehension. Another question needs to be raised at this point. How valid are the SLEP scores that examine students' listening and reading comprehension, when what really needs to be tested here is their speaking and writing skills? Furthermore, the appropriateness of tests that examine language or performance ability during a very short period of time and without taking into consideration external factors that could influence in a negative way the results could be questioned. The author, himself, recognised the need for a more "communicatively oriented test" (Warschauer, 1996). His doubt, though, about the validity of the SLEP test didn't prevent him from gathering his data based on this instrument for data collection.

To collect data during his experiment, the researcher used two methods, observation and monitoring. Each of the four groups participated in two fifteen-minute discussions. The whole experiment lasted 75-minute class period. Two topics of

---

[6] This brings in mind the "Hawthorne effect" (Brown and Dowling, 1998).

discussion were selected so that the experiment was "counter balanced". They were drawn from the current course theme, "the family", in cooperation with the teacher.

During face-to-face discussion students were recorded with a tape-recorder and observed by an outside observer, who made notes about who said what. During electronic discussion the students, facing the walls due to the position of the computers, used the Daedalus Interchange program with which they were already familiar. Their own teacher monitored the electronic discussion, which was usual in this type of discussion.

Several issues need to be raised at this point. The author observed that many students might have been discouraged in participating in the face-to-face discussion, and especially the shy ones[7]. The presence of an external observer and of a tape recorder could have been thought as invaders. Brown and Dowling (1998) refer to this as the "epistemological paradox". This term has been attributed to a situation within a survey, where the observer influences the field of his observation by his/her very presence.

On the other hand, during electronic discussion students were in a very familiar environment in terms of software and procedure, as well as in terms of observation. In addition, since face-to-face discussion was overt, whereas electronic discussion was covert, it can be assumed that in the former case, there might have been some effects that could not be controlled. Therefore, to an extent, it seems that it is very difficult to make comparisons from the data collected under such circumstances.

3.6) EMPIRICAL FINDINGS

The author's empirical findings were related to both concept variables, equality of participation, and quality of the language used in terms of complexity and formality.

To examine the equality of participation, Warschauer made a comparison between the four groups, by calculating the percentage of participation (IPC) for each student, both

---

[7] Could this be the reason why four students were absent at that day?

during face-to-face and electronic discussion. He found that three out of four groups showed greater equality of participation in electronic discussion.

By calculating the Gini coefficient, he found that electronic discussion was generally twice as equal as face-to-face discussion. The Gini coefficient, according to the author, "sums, over all the group members, the deviations of each from equal participation, normalized by the maximum possible value of this deviation" (Warschauer, 1996). It is not though very clear in which level the calculation of the Gini coefficient enables the researcher to make claims about the characteristics of the two forms of interaction.

He also analysed correlations between age, time in U.S., SLEP listening score, SLEP reading score and IPC and provided no further analysis than identifying the correlation between SLEP listening score and IPC as being the largest.

Additionally, the author compared nationality to IPC and found that among Chinese, Japanese, Vietnamese and Filipino students only the last showed decreased participation in electronic discussion, whereas all the others increased their participation during electronic discussion.

Finally, according to the researcher, students' attitude was generally better towards electronic discussion than towards face-to-face discussion.

As stated by the author, none of the aforementioned findings were checked for significance, since it is not really meaningful to carry out tests of significance with such a small sample. This, however, questions the appropriateness of the quantitative comparisons made and of the conclusions drawn in Warschauer's paper, since the lack of any statistical significance raises doubts about the extent to which the numerical differences are due to chance.

To examine the quality of the use of language in terms of complexity, the author calculated the way in which language was used both lexically and syntactically in electronic and face-to-face discussion and found that students' use of more complex language, in both ways, is greater in electronic than in face-to-face discussion.

Another observation made by the author is that students changed turns more rapidly in face-to-face discussion, during which their speech was more direct and comprised of more interactional features, such as questioning, recasting, confirmation checks and paraphrasing than in electronic discussion.

Finally, students tended to use more formal language during electronic discussion than during face-to-face discussion, according to Warschauer.

In general the findings that were analysed quantitatively in relation to the language complexity (TTR, CI) were checked for significance. Further discussion about their validity though is needed and will be made in a further stage of this critical review. The qualitative analysis, which provided the researcher with information about the language formality, was very limited and not explicit at all. Therefore, its validity, as will be examined later, is considered to be doubtful.

4. GENERALISING

4.1) HOW IS THE LINK BETWEEN THE PROBLEM AND THE FINDINGS ESTABLISHED?

At this section of the critical review, an effort will be made to examine how and if the author has managed to link his findings to the concept variables that constitute his problem.

The concept variable "equality of participation" was operationalised and through quantitative measure (IPC score, Gini coefficient), the former was calculated. This was achieved by counting the number of words per speaker and computing the participation percentage for each student in both environments, face-to-face and electronic discussion. The author named this participation percentage, the IPC score[8] (Increased Participation in Computer mode) and based on this score, he compared the participation percentage between groups. The author computed the Gini coefficient of participation inequality for each group of students, based on the IPC score as well.

However, the validity of this measure needs to be examined. Is the number of words an indicator for the amount of participation? In oral speech, body language and facial expressions may substitute many words or whole phrases that could not be avoided in written speech. For example, from the extract provided in the article as typical face-to-face exchange, S2 says: "I live with my parents but…", S3 answers: "Oh, yeah" and S4 answers: "You are independent". S2 stops his phrase and doesn't provide any further information to the reader. However, S3 and S4 seem to understand what S2 wants to say and the reader can assume that S3 and S4 have been able to interpret S2's body language or facial expression. Had the same student used electronic discussion as a medium, he wouldn't have been able to stop mid-sentence and hope to be understood and therefore he would have got a higher IPC score for the same amount of participation.

---

[8] The IPC score was calculated by subtracting students' face-to-face participation percentages from their electronic discussion participation percentages

To calculate the other concept variable "quality of the language used" in terms of complexity, the author used two indicator variables, the Type-Token Ratio (TTR), which is defined by the author as "the total number of different words divided by the total number of words" (1996), and the Coordination Index (CI) defined as "the number of independent clause coordinations divided by the total number of combined clauses" (1996).

Although the findings related to language complexity were checked for significance, there are some questions that need to be raised. Does the total number of different words divided by the total number of words enable the researcher to calculate language complexity? From the extracts provided in the article as paradigms of electronic exchanges, I tried to calculate the TTR for S2 and S3[9]. S3's extract was, could be argued to have been, more structured and more complex than S2's. Yet, they both attained the same TTR score. S3 also got a lower participation percentage than S2, since the extract provided, although meaningful, was shorter than S2's.

It could be argued, at this point, that the TTR is only helpful in comparing the language complexity between face-to-face (oral speech) and electronic discussion (written speech), where a more significant difference of language use can be assumed.

It is worth mentioning at this point the question that some students of the Penn State University addressed to Mark Warschauer, through electronic discussion-mail[10], about this article: "In our various experiments with our chatrooms we have found that language on the contrary often breaks down to a very superficial, abbreviated and sometimes telegraphic style" (Warschauer, 1999). The researcher's response was that several interfaces such as ICQ, IRC etc. encourage short and abbreviated style in written speech whereas "Daedalus Interchange" used by the author "encourages

---

[9] S2: "Most of the people believe housechore is the type of female, but nowaday since male and female are equal. If the husband and wife are full time workers they should share the housechorse. If the wife cook then the husband do the dishes. Or perhaps the husband could help the wife the prepared for dinner. In order words, both husband and wife have to shared the housechore. Men will be able to do everything that women do. For example, clean, cook, wash, take care children. **TTR: 55/85=, 647059.**

  S3: "Based on my experience, the husband do share some of the housework with his wife. Once in a while, he must cook for the entire family and do the cleaning of the house. The husband should help in taking care of the children, like taking the child to the doctor when the wife is working. Also, the husband should do once in a while the groceries. **TTR: 42/65=, 646154.**

[10] http://beetle.la.psu.edu/tifle_spring1999/mark_Q&A.html

people to write whole paragraphs and take a little more time in writing" (Warschauer, 1999). In his answer the author has introduced another variable that influences electronic discussion. The total absence of this variable in the paper can strongly question the validity of the conclusions drawn.

The author made a qualitative analysis to indicate two other differences in the use of language between the two groups, "one related to turn-taking and one related to formality" (Warschauer, 1996). However this qualitative analysis was very limited. The author provided extracts of the transcripts in his article and claimed that they were representative of the discussions monitored both during face-to-face and electronic discussion. No further information has been provided on how the author has analysed his data and on why the extracts provided in the article are representative. Although there is a relation between the qualitative analysis and the concept variable of this study, the lack of "internal explicitness and coherence of the theory" (Brown, A., and Dowling, P., 1998) seems to threaten its validity.

4.2) HOW IS THE RESEARCH GENERALISED?

At the starting point in his article, Warschauer expressed his intention to relate various attributes such as gender, nationality, age, attitude and time in the U.S. to his concept variables. This aspect of the study is very weak. According to the author, the gender variable could not be explored within this sample, not only because four male students were absent but also because there were only six male students in total. Therefore any findings in relation to this variable wouldn't be generalisable.

The same doubt could be expressed about other comparison between Warschauer's concept variables and the various attributes. Can any generalisations be made about Filipino students, out of five Filipinos that participated in the experiment?

The author seems to have a positive answer to the above question and as a consequence he was driven into generalisations in terms of nationality. In more details, he argued that Filipinos were more active during face-to-face discussion

because of their greater practice in speaking English in the Philippines, since English is one of their national languages.

Japanese students, in contrast, not only don't have the same opportunities to practice English in Japan, but they also have cultural differences to Filipinos in that they are socialised in a different way. It is because of this that the researcher finds their restricted participation in oral discussion understandable. One of the Japanese students was surprisingly active in face-to-face discussion. This phenomenon was explained by the researcher as an outcome of the student's marital status. She was married to an American. From these observations Warschauer made the following generalisation: "lack of oral fluency and discomfort in speaking out are important factors in determining students' relative participation in face-to-face and electronic discussion" (Warschauer, 1996).

Although the researcher recognised the small and heterogeneous nature of his sample, he used some members of it not only to represent nationality types but also cultural ones. Is it possible to make sense of the differences between the groups out of such a small sample? This is a question that the researcher should have addressed to himself before choosing this specific sample to examine his problem.

The author made the generalisation that other factors, such as shyness, affect the level of students' participation when trying to explain unexpected findings, from the correlation between SLEP listening score and IPC. Although logical, this still remains a hypothesis.

The nature of the sample produces problems for the main focus of the study as well. The researcher generalised his findings about the equalising role of electronic discussion. Had the sample only contained Filipinos, would it have been possible that the findings would have shown no significant difference in equality of participation between face-to-face and electronic discussion but rather that electronic discussion discourages more verbal students from participating? Alternatively, had another interface like ICQ been used, would the findings have enabled the author to make the same generalisations?

5. CONCLUSION

The researcher has attempted to indicate the difference of language use both in level and in quality, among written and oral speech. The equality of participation and quality of the language use were identified as the concept variables. Both the face-to-face and the electronic environment, as a medium for discussion have been identified as the independent variable in Warschauer's study. The author has also been interested in relating his findings to various attributes, such as gender, nationality, age, language ability, student attitude and time in the U.S.

Interpreting his findings, the researcher made several generalisations. However, there are limitations, which question both the findings and their interpretation. Firstly, the sample of the study was not representative. Although this is very common in educational research, it does not enable the researcher to make generalisations. Secondly, the appropriateness of the techniques used by the author for data collection has been challenged. In several part of the critical review it has been argued that generalisations made in the study are not justifiable from the research findings.

Furthermore, Warschauer's claim of conducting a controlled experiment has also been challenged and the need for more data available to support the data analysis was expressed. Although the restricted amount of data available could be justified due to the space limit that authors are expected to respect, to have their work published, it is worth mentioning that there is no other more detailed version of this article[11].

The author clearly addressed this article to second language teachers. When the author was asked (see footnote 9) whether he had any suggestions on how to deal with teachers who are not willing to implement technology in their classroom, he responded: "We have to find ways to offer positive examples and support over a long-term process" (Warschauer, 1999). Is this study aiming at offering a positive example to teachers who are still not convinced about the benefits of implementing technology in their teaching?

---

[11] The author of this article through personal communication provided this information.

Even though the researcher mainly addressed his study to second language teachers, he acknowledged his readership included researchers as well, since he invited them to read his study and undertake their own within the same field.

The author, in the course of his article, mentioned several weaknesses of his study. He also indicated several gaps in the field and invited other researchers to fill them. These are positive aspects of this research. However, the large number of weaknesses within this study indicates that many aspects of it should be carefully revised.

<u>BIBLIOGRAPHY</u>

Bell, J. (1999). *Doing Your Research Project: A Guide for First-Time Researchers in Education and Social Science* (3rd Edition), Milton Keynes: Open University Press.

Brown, A. J. and Dowling, P. C. (1998). *Doing Research/Reading Research: A Mode of Interrogation for Education,* London: Falmer Press.

Cohen, L., Manion, L. and Morrison, K. (2000). *Research methods in Education* (5th Edition), New York and London: Routledge Falmer.

Warschauer, M. (1999). *Virtual Guest: Mark Warschauer* [on line]. Available at: **http://beetle.la.psu.edu/tifle_spring1999/Mark__Q&A.html**. Last accessed 5th December 2000.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO journal* 13(2), 7-26.