Protecting Expression in Teleconferencing: Pseudonym-Based Peer Review Journals[1]

David S. Stodolsky

Abstract: The social environments of educational systems are less than ideal because power differentials exist that can suppress the free exchange of ideas. One solution is to strengthen personal integrity with an anonymity shield. Many text-based conferencing systems permit anonymous contributions, but this often leads to irresponsible behavior. If people are limited to one and only one pseudonym, however, responsible behavior can be expected. This reputation preserving anonymity overcomes the problems with traditional systems. A reputation is developed through peer evaluation which is based on routinely elicited judgments. Evaluative judgments of a message by one person can be available to all other potential receivers of that message immediately. Evaluations can then be used to automatically select messages worth reading. This approach deals effectively with the problems of both information overload and irresponsible behavior while providing the highest possible protection of expression,

'He told the truth and politicians and civil servants hated him for it.'
(Wright, 1987, p. 356)

The notion of an ideal speech situation has alonghistory. Both Plato in the *Phaedrus* and Habermas argue that if dialogue is-to lead to truth, an ideal speech situation is needed, which in turn presupposes an ideal social environment (Bernstein, 1978, p. 262; McCarthy, 1978). The social environments of educational systems are less then ideal because power differentials exist between students and teachers, and among workers within the educational establishment. One problem is that people may be punished if they speak in opposition to power holders. More often, persons censor themselves because they fear that speaking out will damage their career possibilities.

Anonymity: Potential and Problems

One solution is to strengthen personal integrity with an anonymity shield. Karabenick (1987) has shown that this can have direct educational benefits, since people are more likely to seek help in an academic situation if they can do so anonymously. He argues, "Seeking help [which must be preceded by admission of inadequacy] when needed is an integral part of the learning process" (p. 69). Further, "in cases where anonymity (e.g., using pen names) is permitted the reason [for increased help-seeking] is obviously the reduction in social stigma, or embarrassment.. .especially when faculty have access to the conference" (p. 72).

More generally, it can be argued, based on operant theories of learning, that feedback from the environment is crucial for learning. High rates of response -that create learning opportunities – can be sustained when correct responses are rewarded and erroneous responses are not punished. This requires an environment in which the individual is protected from negative consequences resulting from errors made during the learning process. That is, the protected learning environment acts as a discriminative stimulus for production ofresponses that mayyieldrewards. These rewards can then shape the behavior in a desired direction.

Many text-based[2] conferencing systems permit anonymous contributions. This often leads to irresponsible behavior (Wilkerson, 1987). "Anonymity breeds irresponsibility" (Spragge, 1987, p. 98), was the conclusion of one system administrator. Abuse of anonymity has also been a long standing problem with scientific journals (Garfield, 1988). The current peer review system has been described as one in which power relations have become dominant (Michel, 1982). Inappropriately used, anonymity can generate power differentials as well as irresponsible behavior (Garfield: 1988). Another important limitation of protection based upon anonymity is the inability to reward persons individually for specific acts. This can be seen as a block to effective learning and to motivation. It is a recognized problem with scientific peer review and surely plays a role in academic learning when anonymity is employed.

Pseudonymity

What is needed is reputation preserving anonymity, or pseudonymity, which overcomes many of the problems with traditional journals while ensuring individual integrity. If people are limited to one and only one psuedonym, responsible behavior can be expected. The person must protect their pseudonym from developing a bad reputation or others will not select messages or use judgments issued under that pseudonym. On the other hand, messages under a pseudonym with a good reputation will be read more often and judgments as to what is worth reading under that pseudonym will have a strong influence on dissemination of messages.

An often raised objection to this approach is that even if an author can not be traced through the system because of formal protections, the writing style and similar factors can be used to identify an author. In the system described here, judgments similar to votes play an important role, and these can provide complete protection even in relatively small groups. The system automatically collects judgments on messages after they are read and transmits them in a standard format. Completely protected judgments can then be used by others to automatically select messages worth reading. We assume that it is not feasible to read all messages.

Public versus Protected Communications

An analogy between the proposed system and anonymous voting in parliamentary meetings can be extended, although it is not completely correct. In such meetings, communication occurs on two levels, public and anonymous. Substantive motions, procedural motions, and discussions occur on the public level: There is no attempt to hide source identity. Voting, however, is typically anonymous. Similarly, with the proposed system substantive messages and judgments directed against them can be considered as operating on two different levels. Especially in a smaller group, it may be difficult to effectively hide the source of a substantive message because of stylistic features. Judgments in a standardized form, however, have no such features and therefore offer significantly greater protection for author identity. Like anonymous votes, they can play a crucial role in decision making.

However, unlike anonymous votes, pseudonymous judgments can play an important role in reputation development. For instance, a person could develop a good reputation purely on the basis ofjudgmental responses: those responses that are highly protected. This would ensure that any substantive messages later contributed would immediately come to the attention of other group members, thereby maximizing the message's influence potential. With the pseudonym system there is also much greater flexibility in the treatment of substantive messages. One extreme could be to have them signed by their authors using their public names. Another would be to have substantive messages sent first to human or machine editors, who would remove stylistic features that could reveal authorship, translate them to another language or even rewrite them for clarity. Pseudonymous communication, then, offers many more possibilities for finding an effective balance between protection and accountability for authors. It also offers significantly more information to readers – actually their programmed message sorting systems, – who must decide which messages to read.

Basic Organization of the System

In its simplest form, the system includes untraceable mail and digital signature capabilities. By untraceable mail we mean that a message can not be traced back to its sender by physical means or by analysis of the information transmitted with the message. An ideal broadcast system would have such a physical characteristic. In practical systems, a ring topology network can transmit untraceable mail at 25% efficiency as compared to normal mail. The author's identity can be unconditionally secure, that is, resistant to infinite computational power. Then, finding the source of the message requires cooperation of all parties except for the one being traced (Chaum, 1985). A less secure, but readily available system is the public-access telephone network. Many data networks can also provide adequate security for short connection times.

The digital pseudonym[3] is required to be untraceable and unforgeable. A one-to-one mapping between persons and pseudonyms is required. This can be implemented withpublic-key *cryptography** using an independent registrar or is-a-person organization. This organization is the only one that can engage in pseudonym creation. An interaction with a potential user permits the authorization for creation of a pseudonym to be issued to the user. The user then, at a later time, returns by untraceable mail the actual pseudonym to be used. This pseudonym serves as that individual's public key in a digital signature system. Messages decrypted with that key could only have been sent by that individual. Persons must be physically identified to obtain an authorization, thus each person can acquire one and only one unforgeable pseudonym.

One objective of the pseudonym system is to focus a reader's attention as completely as possible on the content of a message. The pseudonym mechanism makes it very difficult to determine an author's identity, thereby discouraging giving any attention to this aspect of a message. In fact, the pseudonyms discussed here would not be in a form easily read or remembered by a reader, normally they would not be seen at all. Readers would train their computer systems as to which pseudonyms merited attention merely by giving evaluative responses to messages. Both the reputation of the author of a message and the reputations of previous readers of that message (assuming they offered judgmental responses to it) would be used to automatically rank the message in priority. This is meant to duplicate, in a more rigorous manner, the way we use recommendations of friends and colleagues to decide what is worth reading.

A common misconception about using pseudonyms is that the benefits would be short-lived since once a pseudonym's reputation had been established it would function just as a real name, prejudicing reader acceptance of massages and reactions to them. It is true that only an anonymous message system guarantees the evaluation ofmessage content without any influence of previous messages from that author. However, such a system also offers the reader no basis for selecting messages to be read. In a properly functioning pseudonym-based system, reputation information gives an unbiased estimate ofwhether anew message is worth reading. This estimate isbasedpartlyupon the content of previous messages from the same author. The author's institutional position, the prestige of the author's institution, and other biasing factors which typically influence readers are screened out. If these factors were true indicators of message quality, then they would be correlated with reputations developed within the message system. In summary, a system based upon anonymous messages treats each message equally, while a system based upon pseudonymous messages treats each author equally.

Differential Competence and Reputation Management

The possibility of differing levels of competence in different subject areas can be accountedforby allowing persons to have a different pseudonym in each separate conference or journal. By use of a *credential mechanism* (Chaum, 1985), expertise developed in one conference can be transferred to another without any loss ofsecurity- that is, without release ofinformation thatwould permit the association of different pseudonyms. If one of the names used was the name by which the person was known to an educational institution, this mechanism could be used to show that educational objectives had been satisfied.

Similarly, the mechanism could permit reputational credentials developed outside the message system to be moved, untraceable, into the system. Thus, for example, a given message could be shown to be from someone who had received a certain educational degree, to have achieved a certain academic rank, or to be employed by a certain institution. The more detailed such information was, however, the more constrained the set ofpossible authors and the more limited the protection for those authors.

An essential feature of the credential mechanism is the ability to move reputational information from one name to another untraceable. Given this ability, even positive identification of the author of a given message would not compromise the overall functioning of the system, since all pseudonyms could be changed without a loss of reputational information. For instance, such a procedure could operate very similarly to the double-blind peer review used by many scientific journals. While a message was being evaluated, authorship could remain hidden. Later, the author could claim the message and even the referees could identify themselves publicly. If this was followed by an immediate change of pseudonyms, then the next message from that same author could be evaluated in an unbiased manner, since the association between the message with the publicly identified author and the new message would be untraceable. The credit for producing that publicly identified message would be available to the author, however, thus ensuring that the new message was widely distributed.

In an educational setting the exchange of pseudonyms might be necessary if teachers or staff members were to retain their protection. Otherwise, the continuing presence of their pseudonyms in a conference, while student names were constantly changing could give away their identity. Also, there are instances in which a person might wish to share a program or text file that would reveal the author's identity That is, the work in question might have a known author or be found in a storage location that belonged to a single person. However, once pseudonymous communication became the dominant mode of interaction, expertise established under a pseudonym would be connected to specific works also available under that name, thus the previously mentioned situation would not require a change of pseudonyms.

Often it is helpful to have publicly known local experts available in order to get quick answers to specialized questions. In a conferencing system environment the element of geographical locality is eliminated. It is replaced by content or subject matter locality permitting consultation with the leading expert on the specific topic. Identification of such experts would be, at least in part, automatic as a result of the system for reputation maintenance discussed here.

Review Messages

When an individual reads a message and makes a judgment of it, that judgment can be signed by the reader and broadcast to other potential readers. These judgments of a message can be used by those who have not yet read a message to rank it in priority. The standard format of the review permits the user to allow a program to compute the probable importance of a given message and automatically schedule messages for attention. A Bayesian estimation model can be used to combine the information about the author with the judgments of the previous readers. The user's own judgment upon reading a message can then be used as a basis for revising the probabilities in the model parameters that describe each person's judgmental competence and competence as an author.

Evaluative Dimensions

Another elaboration of the basic system permits judgments to be given on multiple dimensions. These judgments establish different types of formal relations between messages. For instance, a scientific paper after having been judged *relevant* is most likely to be accepted for publication if it meets three criteria (Garfield, 1988). First, it should be *sound.* The author must have employed reliable data, drawn valid conclusions, and committed no flaws in logic. It should also be *original.* Finally it should be *significant,* meaning that it should contain some new perspective or observation of potential importance. Judgments on these dimensions could be combined to decide whether a message is worth reading.

While evaluative messages could give quantitative responses on various dimensions, this is not a requirement for system operation. Both the dimensions or categories for evaluation and the scaling of such evaluations would follow from agreements between the users of the system. Such agreements would permit more or less effective sorting of messages by computer software. With very high message volume, multi-dimensional and carefully scaled responses would be beneficial. However, if powerful natural language understanding software was available, then unstructured responses could be entirely adequate.

The need for evaluative information becomes much greater with the complex and opaque multi-media or hypertext documents now being developed (Stodolsky, 1987). With parts of such documents or with very short messages and more conversational interaction, often associated with voice messages, the types of judgments could be quite different (Stodolsky, 1984). The degree of impact on the priority relations among messages waiting to be read would be the crucial measure of quality for review messages.

If a reader finds a message to be lacking on a certain dimension, a substantive message may be offered to supplement the judgment given. A structured form of argumentation can then take place. Some authors would be attracted into the controversy and strive to gain credibility by issuing judgments referencing a given message. Others would prefer to wait until the situation had stabilized, as calculated by Bayesian estimation, before reading any of the messages (Stodolsky, 1984). This latter strategy might be called the text or reference book approach to conferencing. The first might be called the meeting approach.

Summary of the Approach

The security of the system reduces the effect of power relations on the interchange of information (Stodolsky, 1985). The judgment mechanism focuses attention on reasoning in the dialogue process. The formal relations among messages and quest for credibility attracts competent criticism. The overall system integrates the reliability of the scientific journal with the rapid response of informal dialogue, thereby creating a powerful educational technology.

IMPLEMENTATION

A fully developed system of the type outlined here would be quite adequate to support a multiple journal publication program of a major scientific society. In fact, it would be an improvement over the best current practice. Significant educational objectives can be supported with much less elaborate procedures particularly if the demands for security, flexibility in registering of new participants, and transferability of credits are relaxed. This is appropriate in an educational setting where the intensity of assets is significantly reduced as compared to the professional environment. That is, the size of payoffs and therefore, motivations toward corruption are reduced in educational environ: ments.

A *core* mechanism of the proposed teleconferencing system is pseudonym-based communication. While this can be implemented with the highly secure cryptographic techniques mentioned, it can also be accomplished with much simpler procedures in educational settings. For example, if a third party can be found who is trusted by all participants, then that person can simply be assigned the responsibility of seeing that each person receives one and only one pseudonym, that is, act as registrar. If that person also plays the role of computer system administrator, then standard controls and accounting procedures available with current conferencing systems can also be used. What is crucial is that the users of the system feel they are secure, so that educational benefits of pseudonymous communication can be attained.

When this level of security is inadequate, another procedure is available assuming participants can meet physically at the beginning of an educational program. Then the number of persons present can be determined and that exact number of paper slips with pseudonyms and computer passwords written on them can be placed in a hat. The hat is then passed with instructions that each person select one and only one slip. If this procedure succeeds that is, if each person gets a valid pseudonym and password, then protected communication can occur on any conferencing system. The validity of the slips could be checked, in the first instance, by a signature or other unforgeable mark on the slips. A further validity check and security enhancement would occur when participants logged-in to the computer for the first time and changed their passwords. This would best be accomplished before participants dispersed. A second "log-in" to ensure that the new passwords were functioning correctly would avoid possible problems that commonly occur when inexperienced persons begin using computer systems.

While ad-hoc procedures such as this can give an adequate level of security, they have their limitations. Since pseudonyms and their computer accounts will most likely be terminated at the conclusion of the educational program, and since credits earned under these names cannot be transferred or preserved (unless protection is compromised), it is likely that concern with maintaining good reputation ofpseudonyms will drop toward the end of the program. Under such conditions an increase in irresponsible behavior and in the number of abusive messages can be expected.

System Specification and Software Development

A level of implementation complexity beyond the ad-hoc arrangements mentioned above assumes enhancements to software. The process can be divided into system specification and software development. Two packages of software and associated institutional arrangements would be included in specification of a complete system. First, a powerful communications software package, that includes a display interface manager, communications handler, reputation database, statistical estimation routines, mail sorting software, and cryptography subsystem would be specified. Second, organization of the registrar function and associated cryptographic security system would be specified. The independence and security of this function is critical to the success of the pseudonym system.

From afunctional point of view, software development can proceed in three lines. First, a mechanism for selecting messages based upon both content and source can be developed. The SMART current awareness information system (Fox, could be used for content-based selection. For reputation-based selection, a reputation database, statistical estimation routines, and mail sorting software would be integrated. Selection using both content and source would require integration of the two selection mechanisms with a display interface manager and communications handler.

Second, the collection of evaluative responses is a relatively straight forward software development and human interface problem. Unless the software integration is very smooth and transparent, however, is unlikely that readers will use the facility. The incentives for contributing evaluations of messages is likely to develop when substantial volumes of material with conflicting points of view are exchanged.

A third line of software development includes the security mechanisms for constructing pseudonyms. As the incentives for contributing messages increased, the need for protection and authentication would also increase. Substantial development work is now in progress to replace the current password-based authentication on computer systems with public-key based cryptographic authentication procedures. These authentication procedures use the same principles as the pseudonym mechanism needed for the proposed system. At least one public-key cryptographic system is currently available as a commercial product.

Implementation Strategy

These procedures assume that physical tracing of communication is not practical. On most computer systems it is quite easy to determine which participant is using a certain line into the computer. In the case of dial-up lines, however, there are significant legal as well as technical barriers against determining actual identities (Stoll, 1989). Dependable security, however, requires specially structured communication subsystems (Chaum, 1985).

The focus of efforts in implementation can be guided by the actual needs in a given educational environment. When a high volume of messages on different subjects is expected, attention might best be focused upon an adaptive sorting mechanism based upon topic categories. In the case of a high volume of messages of variable quality, the source evaluation mechanism would be very important. This peer reviewing of messages might also be seen as desirable because it is an effective means of providing feedback to learners. It can also serve as a tool to evaluate relative competence of participants assuming an adequate level of competence in the group as a whole. When power relations threaten the free exchange of statements and judgments giving adequate attention to the security questions could be crucial. Effective sorting of messages is limited by the quality of judgments directed to them The quality of judgments is in turn dependent upon the degree to which the speech environment approaches the ideal of non-dominative communication.

NOTES

[1] Earlier versions of this paper were presented August 11, 1988, at the Fourteenth World Conference on Distance Education, Oslo, Norway and electronically prepublished May 9, 1989 in *Communication Research and Theory Network*, No. 775. University Park, PA: The Pennsylvania State University, Department of Speech Communication.

[2] *Text-based* or asynchronous (store-and-forward) conferencing systems are distinguished from real-time audio conferencing which requires simultaneous presence.

[3] The *pseudonym* is a binary number of about two hundred digits.

cryptography uses two different keys, one for encoding and one for decoding. The public-key can be widely distributed without risk of revealing the private-key that is used to decode messages and sign documents.

This system makes key distribution practical when there are large numbers of users.

REFERENCES

Bernstein, R.J. (1978). *The reconstruction of social and political theory.* Philadelphia, PA: University of Pennsylvania Press.

Chaum, D. (1985). Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM, 28, 1030-1044.*

Fox, E. (1981). Implementing SMART for minicomputers via relational processing with abstract data types. *Sigsmall Newsletter,* 7(2), 119-129.

Garfield, E. (1988). Refereeing and peer review: Part one. In E. Garfield, *Essays ofan information scientist: Towards scientography* (Vol. 9, pp. 230-238). Philadelphia, PA: ISI Press,

Karabenick, S.A. (1987). Computer conferencing: Its impact on academic help-seeking. *The second Guelph Symposium on computer conferencing,* (pp. 69-76). Ontario, Canada: Guelph University, Department of Rural Extension Studies.

McCarthy, T. (1978). *The critical theory of Jurgen Habermas.* London: Hutchinson.

Michel, F.C. (1982). Solving the problem of refereeing. *Physics Today,* 35(12), 9;82.

Spragge, J.G. (1987). If computer users are a community, the conference must be the town meeting. *The second Guelph symposium on computer conferencing,* (pp. 91-104). Ontario, Canada: Guelph University, Department of Rural Extension Studies.

Stodolsky, D. (1984, December). *Self-management of criticism in dialog: Dynamic regulation through automatic mediation.* Paper presented at the symposium Communicating and Contracts between people in the Computerized Society, Gothenburg University, Sweden.

Stodolsky, D. (1985, June). Telematic journals and organizational control: Integrity, authority, and self-regulation. *Abstracts of the 7th colloquium of the European Group for Organization Studies,* Saltsjoebaden, Sweden.

Stodolsky, D. (1987, April). Telematic journals [Abstract]. *Proceedings of the First STIMDI Conference on Man-computer Interaction.* Stockholm, Sweden: STIMDI (Sveriges tvaervetenskapliga intressefoerening foer maenniska-datorinteraktion).

Stoll, C. (1989). Stalkingthewilyhacker. *Communicationsofthe ACM, 31,484-497.*

Wilkerson, I. (1987, April 18). Ethnic jokes in campus computer prompt debate. *The New York Times.*

Wright, P., with Greengrass, P. (1987). *Spycatcher.* Richmond, Victoria: William Heinemann Australia.

AUTHOR

David S. Stodolsky is a Guest Researcher in the Department of Psychology Copenhagen University, Njalsgade 88, DK-2300 Copenhagen S, Denmark:

ACKNOWLEDGEMENT: The author gratefully acknowledges the assistance of the following persons for their comments on drafts of this article: Bill Hunter Kathleen M. Carley, Terry Winograd, Gerald Phillips, Hedwig Miller, George Altomare, and two anonymous reviewers.