Japanese Psychological Research
2002, Volume 44, No. 3, 173–179
Special Issue: Reciprocity between applied and basic studies

Why are four eyes better than two? Effects of collaboration on the detection of errors in proofreading

YOSHIAKI NIHEI
Department of Psychology, Graduate School of Arts and Letters, Tohoku University, Sendai, Miyagi 980-8576, Japan

MASAYOSHI TERASHIMA
Department of Psychology, Faculty of Arts and Letters, Tohoku University, Sendai, Miyagi 980-8576, Japan

IKUO SUZUKI
Department of Psychology, Graduate School of Arts and Letters, Tohoku University, Kawauchi, Miyagi 980-8576, Japan

SHINICHI MORIKAWA
Institute of Social Research, Institute of Nuclear Safety System, Mikata-gun, Fukui 919-1205, Japan

Abstract: We investigated why collaboration has positive effects on performance in detecting errors. Forty-eight undergraduate students worked on a proofreading task individually or in pairs. In the first group, 16 subjects performed the task individually (the individual group). In a second group, 16 pairs of subjects performed the task in dyads, discussing the errors detected (the collaborating group). The results show that the individual group detected significantly more "contextual errors," which are found by tracking the context of a passage, than "surface errors," which are based on the meaning of a word. In contrast, the collaborating group detected significantly more surface errors than contextual errors. The surface error detection rate of the dyads was high enough that it could not be explained by a simple interpretation such as two additional eyes found errors missed by the first two eyes. The results indicate that proofreading in dyads is beneficial because it enables the subjects to free themselves from a default style of reading. Moreover, for the collaborating group, the rate of error detection tended to increase in the latter half of the work, suggesting that interaction between the two individuals gave them an opportunity to learn about the task by cognitive restructuring.

Key words: collaboration, proofreading, human errors, learning.

Working in dyads is often recommended in inspection tasks where safety is of particular concern. As to the performance of dyadic works in general, researchers have reported two contradictory results. On one hand, investigators have found that the collaboration of two individuals produced higher levels of performance than when they both worked individually; that is, the work of the pair was more than the sum of the work of each individual (e.g. Teasley, 1995). On the other hand, investigators have found that the performance per head in dyadic works was lower than that in individual works (e.g. Andersson & Rönnberg, 1995).

These contradictory results have been explained from two major theoretical perspectives:

motivational and cognitive theories. In motivational theories, some authors (e.g. Harkins, 1987) explain the superior performance by dyads in terms of social facilitation (Zajong, 1965). Other authors (e.g. Comer, 1995) interpreted the inferior performance by dyads in terms of social loafing (Latanè, Williams & Harkins, 1979). Thus, some books on industrial safety and human error claimed that human "double-checking" is sometimes hazardous in inspection works (e.g. Hashimoto, 1988; Senders & Moray, 1991; Safety Research Laboratory, 1998).

Some authors referred also to cognitive theories in order to explain the improvement of collaborative performance as well as its deterioration. Some reasoned that working in dyads is beneficial because pairs of people may exchange their perspectives or cross-cue each other. Thus, exchanging perspectives or cross-cueing causes cognitive restructuring of the two individuals, which in turn improves their performance (Azmitia, 1988; Meudell, Hitch, & Boyle, 1995). This line of reasoning differs fundamentally from a simple explanation, known as "the gleaning theory of error detection," which claims that two additional eyes can find errors that two earlier eyes missed.

In contrast, some researchers have also tried to explain the lower performance per head in terms of cognitive factors (e.g. Wiersema & Van Oudenhoven, 1992). Wiersema and Van Oudenhoven assumed that if two people talk to each other over a task that is too simple to need their reasoning mechanisms, they may be wasting time and the result may be a lower performance.

There may be some reasons for the above contradictory results. The tasks to be performed and the subjects' ages varied across studies: remembering by adults or undergraduate students (Stephenson, Abrams, Wagner, & Wade, 1986; Clark & Stephenson, 1990; Yarmey, 1992; Andersson & Rönnberg, 1995; Andersson & Rönnberg, 1996), proofreading by undergraduate students (Riefer, 1993), spelling achievement by children (Wiersema & Van Oudenhoven, 1992), text learning by undergraduate students (McDonald, Larson & Dansereau, 1985), problem-solving by children (Azmitia, 1988) and so on. The studies varied also in the way of performing the tasks, from a working style in which two individuals worked separately on the same task, to a style in which two individuals were in interaction discussing the task.

In the present study, we investigated the effects of collaboration on proofreading because it is a type of error-detection task that has relevance to safety to some extent. A few studies have already investigated the effects of collaboration on proofreading but these studies were based on "the gleaning theory of error detection" that assumes the effectiveness of "double-checking." They did not use a procedure that allows the subjects to restructure their cognition on the task by exchanging different perspectives between the two individuals.

For example, Riefer (1993) reported an experiment that tested the effectiveness of a proofreading strategy known as team proofreading. Riefer compared the errors detected under conditions where one subject read aloud to a partner who would follow silently, with those under conditions where both subjects read two passages, one aloud and the other silently, but did so separately. The results showed that team proofreading was not effective, because the two readers detected as many errors when they were reading separately as they did when working as a team.

As Sato (1996) pointed out, it may be important in collaborative work or learning that interaction among the subjects produces critical information that leads directly to new learning or comprehension. In Riefer's experiment, team proofreading did not take the form of work where two individuals communicated their ideas about the errors, so it could not be called collaboration as in its fundamental meaning.

For the reasons mentioned above, we compared proofreading by individuals with proofreading where two individuals worked together by freely discussing the errors found, so that interactive communication enables their cognitive change. Under such conditions,

we aimed at investigating the positive effects of working interactively on the detection of errors.

Moreover, the type of errors for detection was introduced into the present experiment as an independent variable. We placed two types of misprints in the passages. One type of misprint was a "surface error" that could be detected based on knowledge of the meaning of a word, and the other type of misprint was a "contextual error" that required tracking of its contextual meaning in order to be detected. The two types of errors varied in terms of the information processing needed during performance of the tasks. We expected that introducing these two types of error into the present experiment would allow us to analyze how the subjects' notion of the kind of errors to be detected was restructured through collaboration.

Method

Design

The subjects were assigned to one of two experimental groups. The experiments were run individually (the individual group) or in dyads (the collaborating group). The type of misprints to be detected and corrected (i.e. the surface or the contextual errors), was a within-subject variable.

Subjects

Forty-eight male and female undergraduate students participated. Their mean age was 21.67 years (*SD* 2.26). Sixteen subjects were allocated to the individual condition, and 32 subjects were allocated to the collaborating condition in 16 same-sex pairs. The ratio of male to female subjects was 7:9 both in the individual and the collaborating groups.

Materials

The subjects read two passages written in Japanese by the first author of the present paper. The first passage, titled "Spoonerism reconsidered," was taken from a non-popular journal. The second passage, titled "Priming plays," was a manuscript prepared as one chapter of a book. None of the subjects had read the passages before.

The two passages were printed in five pages each, so that the total number of pages of the proof was 10, and the number of characters in the proof was 7839. The subjects read the two passages successively.

Two types of misprints were involved in the two passages: surface and contextual errors. The *contextual errors* were those subjects could find only when they read the passages and integrated information from the context or their own knowledge: e.g. doctor (医者) → nurse (看護婦), English (イギリス人) → American (アメリカ人), went (行った) → did not go (行かなかった). The other type of error, *surface errors*, were errors such as omission of letters within a word (e.g., 子どもたち → 子もたち), transposition of letters within a word (e.g., それ以前 → れそ以前), substitution of a Kanji character for a similar one (e.g., 隅然 for 偶然).

The passages had 10 misprints each for contextual and surface errors; thus, the sum of errors was 20. Embedded errors on a page varied from zero to three, and the mean number of errors per page was two. Passages were bound in a booklet, so that the subjects in a collaborating group could not proofread separately. They were asked to use one booklet per pair of subjects.

Procedure

The experiments were carried out in a sound-attenuated room. The subjects proofread individually or in collaborating dyads. The subjects were asked to read the passages in the booklet, to circle the words that they found to be erroneous, and then to write the correct words beside the circled ones. They were given no information on the types of errors, but were only instructed to detect and correct misprints.

Among the collaborating condition, it was emphasized that subjects must not share their tasks, such as one reading the passages and the other checking the text, but should read the passages and discuss the errors detected. A booklet for proofreading was put on the center of a table where two subjects seated side by

side. A video camera recorded the front view of working and conversation states within a dyad.

Results and discussion

Detected errors

The mean working time was 1092.3 s (*SD* 296.7) in the individual condition, and 1656.3 s (*SD* 536.7) in the collaborating condition. The difference was significant between the two conditions ($t(30) = 3.873$, $p < 0.001$). This difference may be reasonable, considering the fact that dyadic work requires time to discuss the errors.

The collaborating group detected a significantly higher number of errors than did the individual group ($t(30) = 3.818$, $p < 0.001$). The former group detected approximately half of the embedded errors ($M = 9.94$, *SD* 3.26), whereas the latter group detected only about one-third ($M = 6.31$, *SD* 1.96).

However, comparing the performance between the two groups for contextual errors and for surface errors separately, we found that the main source of difference between the groups in the number of errors detected was in the difference of surface errors (Figure 1). The collaborating group detected significantly more surface errors than the individual group ($t(30) = 5.286$, $p < 0.001$): 5.562 (*SD* 3.996) for the collaborating group, and 2.438 (*SD* 1.596) for the individual group. For the contextual errors, there was no significant difference between the two groups (4.375 (*SD* 3.317) for the collaborating group, 3.875 (*SD* 2.383) for the individual group; $t(30) = 0.838$, $p = 0.409$).

The individual condition was more advantageous for detecting contextual errors than for detecting surface errors ($t(15) = 2.830$, $p < 0.05$). On the contrary, the collaborating condition allowed subjects to detect more surface errors than contextual errors ($t(15) = 2.366$, $p < 0.05$).

The mean number of surface errors detected under collaborating conditions was more than twice those detected under individual conditions (5.562 vs. 2.438). Had the two individuals worked separately, they may have detected a substantial number of errors in common. In fact, 50% of the surface errors were detected by more than 20% of the individual group subjects. Therefore, the sum of the items detected independently as errors by two individuals might be lower than twice the number of errors detected by dyads.

If we consider the above possibility, the results suggest that collaborative work yields more than the collected work of individuals.

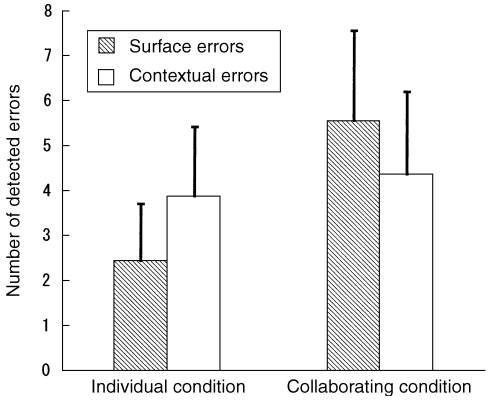

Figure 1. Mean number of detected errors as a function of the type of errors (surface vs. contextual errors) and of the working condition (individual vs. collaborating).

A paradoxical correlation between working time and performance

To understand the difference in the performance of the collaborating and individual groups, we need to know in detail how each subject or dyad worked for proofreading.

Therefore, we calculated the correlations between the number of errors and the time required for work. For the individual group, there was a significant positive correlation between the numbers of detected contextual errors and time used for work ($r = 0.508$, $p < 0.05$). That is, the longer a subject worked, the more contextual errors were detected (Figure 2).

On the contrary, there was an almost significant negative correlation between the number of detected surface errors and working time

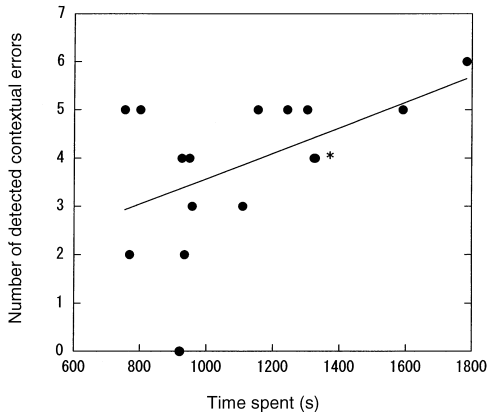

Figure 2. Correlation between the number of detected *contextual errors* and the time spent on the proofreading work. Each dot denotes the result of a subject in the individual condition. The dots with an asterisk(*), which appear to be one dot, are two overlapped dots with very close values.

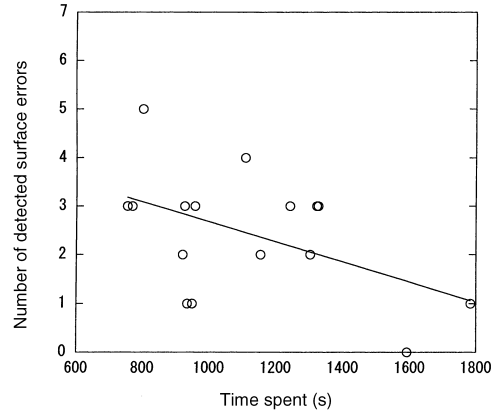

Figure 3. Correlation between the number of detected *surface errors* and the time spent on the proofreading work. Each circle denotes the result of a subject in the individual condition.

($r = -0.486$, $p = 0.056$). That is, the more slowly a subject worked, the fewer surface errors she/he tended to find (Figure 3). This result is paradoxical because it contradicts the conventional speed-accuracy trade-off principle (Fitts, 1954) that predicts a poor performance from faster work.

In the collaborating condition, correlations between working time and the number of detected errors were not significant; $r = 0.377$, $p = 0.206$, for contextual errors, and $r = 0.145$, $p = 0.598$ for surface errors.

We interpreted the results as follows. Generally speaking, it may be a type of "default" style to read while tracking the contextual meaning of a passage. If a subject reads a passage in the default style, surface errors will be difficult to detect using the function of contextual completion (Lindsay & Norman, 1977). The more accurately a reader tries to grasp the meaning of a passage in the default style, the longer time she/he needs to read the passage. Consequently, in the individual condition, a longer working time may have positively correlated with the number of detected contextual errors.

A subject can detect surface errors, in contrast, by local information processing of a word, without tracking the contextual meaning of a passage. If the subjects track the meaning and thus take a longer time to read, they correct surface errors automatically leading to difficulties in detecting surface errors. Thus, we obtained a paradoxical negative correlation between the detected surface errors and working time.

Under the collaborative condition, the correlations disappeared. These results suggest that collaboration worked against a bias, which was to rely on only one of the two reading styles. Under the individual condition, the "default" style was the one in which the subjects track the contextual meaning of the passages. One of the effects of collaboration may be that the exchange of different perspectives allowed them to eliminate the default reading style.

The mechanism of cognitive restructuring

The mechanism of cognitive restructuring through collaboration is considered as follows. In the present experiment, the subjects were only asked to detect and correct errors, but they were not told about the types of errors in the passages. The subjects in the individual group had to discover the kinds of errors on their own.

The subjects in the collaborating group, in contrast, were able to learn about the types of errors, directly from their partners' references, such as "Can we include this wrong character in the errors? Is it a new kind?" In fact, dyads often referred to the types of errors. Their direct references to error types could have a function in reconstructing meta-cognition on the nature of errors.

However, comparing the dyads who referred directly to the error types during proofreading with those who did not, we could not find any significant difference in the number of detected surface or contextual errors. These results might be explained by the fact that the collaborating subjects had another way to learn about the error types. In addition to direct reference by their partners to the types of errors, they could infer the types of errors through the examples of errors that their partners detected. This could be taken as "learning by example," or a kind of "cross-cueing" (Meudell, Hitch, & Boyle, 1995).

Another aspect of collaboration: Effects on learning for further works

Collaboration had another important advantage over individual work: the effect on learning for further work.

Comparing the performance of the individual group in the first half of the work with the latter half, we could not find a significant difference for any type of detected errors ($t(15) = 1.046$, 0.745, 0.155, for surface errors, for contextual errors, and for the total errors, respectively) (Figure 4).

In the collaborating group, however, the total number of detected errors increased almost significantly in the latter half of the work ($t(15) = 1.861$, $p = 0.082$), although the increase in surface or contextual errors did not reach significance ($t(15) = 1.346$, $p = 0.198$ for surface errors, $t(15) = 0.899$, $p = 0.382$ for contextual errors). These results suggest that the collaborating subjects learned the types of errors from interaction in the first half, and took advantage of the knowledge in the second half.

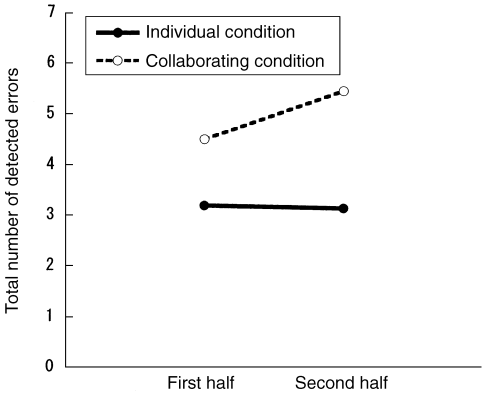

Figure 4. Changes in performance of error detection from the first half to the second half of the proofreading task. The dots and the solid line denote the results from the individual condition, and the circles and the broken line denote the results from the collaborating condition.

General discussion

Collaboration increased the probability of detecting surface errors that tended to evade individual subject's eyes. As mentioned earlier, it has been claimed that human "double-checking" is sometimes hazardous in inspection works. But "double-checking" in the above contexts seems to mean a checking system without interaction between the individuals within a dyad.

When two individuals discussed and exchanged their ideas, as was the case in the present study, the detection rate was superior to the sum of two tasks performed independently. The mechanism of increase was considered to be cognitive restructuring. Both subjects in the pair whose performance was the highest in the collaborating group reported, after their work, that they tended to find types of errors that were different from each other. This kind of complementarity may be an advantage of collaboration.

Collaboration in checking work may be an effective strategy for safety, if two individuals work without social loafing, if they can exchange their perspectives easily, and if they make use of their differences.

Moreover, collaboration improved not only the performance of the work concerned but also the performance in the second half of the work. Experiencing a perspective different from one's own might lead to learning a new perspective, and this could have educational effects that benefit further work. Ames and Murray (1982) indicated that even incorrect information that conflicts with a prior but equally erroneous perspective can stimulate cognitive growth. An interesting question to investigate further is whether the learning effects of collaboration will carry over to the next work that is performed under individual conditions.

Finally, the evaluation of the present investigation into whether four eyes are better than two was based on quantitative norms. It would be worthwhile clarifying whether the advantage of four eyes is maintained in a future study based on qualitative norms.

References

Ames, G. J., & Murray, F. B. (1982). When two wrongs make a right: Promoting cognitive change by social conflict. *Developmental Psychology*, **18**, 894–897.

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: joint recall effects of friendship and task complexity. *Applied Cognitive Psychology*, **9**, 199–211.

Andersson, J., & Rönnberg, J. (1996). Collaboration and memory: effects of dyadic retrieval on different memory tasks. *Applied Cognitive Psychology*, **10**, 171–181.

Azmitia, M. (1988). Peer interaction and problem solving: when are two heads better than one? *Child Development*, **59**, 87–96.

Clark, N. K., & Stephenson, G. M. (1990). Social remembering: quantitative aspects of individual and collaborative remembering by police officers and students. *British Journal of Psychology*, **81**, 73–94.

Comer, D. R. (1995). A model of social loafing in real work groups. *Human Relations*, **48**, 647–667.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, **47**, 381–391.

Harkins, S. G. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology*, **23**, 1–18.

Hashimoto, K. (1988). *Safety Ergonomics*. Tokyo: Japan Industrial Safety and Health Association. (In Japanese).

Latanè B., Williams, K., & Harkins, S. (1979). Many hands make light the work: the causes and consequences of social loafing. *Journal of Personality Social Psychology*, **37**, 822–832.

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: an introduction to psychology* (2nd ed.). New York: Academic Press.

McDonald, B. A., Larson, C. O., & Dansereau, D. F. (1985). Cooperative dyads: impact on text learning and transfer. *Contemporary Educational Psychology*, **10**, 369–377.

Meudell, P. R., Hitch, G. J., & Boyle, M. M. (1995). Collaboration in recall: do pairs of people crosscue each other to produce new memories? *Quarterly Journal of Experimental Psychology*, **48**, 141–152.

Riefer, D. M. (1993). Behavior engineering proposals: 5. An experimental comparison of team versus solo proofreading. *Perceptual and Motor Skills*, **76**, 111–117.

Safety Research Laboratory (1998). *Key points for prevention of accidents from the view point of human factors: cars, facilities, and electrical apparatus*. Tokyo: East Japan Railway Company. (In Japanese.)

Sato, K. (1996). *The world of reading from a viewpoint of cognitive psychology: toward dialogue and cooperative learning*. Kyoto: Kitaohjishobo. (In Japanese.)

Senders, J. W., & Moray, N. P. (1991). *Human error: cause, prediction, and reduction*. New Jersey: Lawrence Erlaum Associates.

Stephenson, G. M., Abrams, D., Wagner, W., & Wade, G. (1986). Patterns in recall: collaborative order in the recall of a police interrogation. *British Journal of Social Psychology*, **25**, 341–343.

Teasley, S. D. (1995). The role of talk in children's peer collaboration. *Developmental Psychology*, **31**, 207–220.

Wiersema, B., & Van Oudenhoven, J. P. (1992). Effects of cooperation on spelling achievement at three age levels (grades 2, 4, and 6). *European Journal of Psychology of Education*, **7**, 95–108.

Yarmey, A. D. (1992). The effects of dyadic discussion on earwitness recall. *Basic and Applied Social Psychology*, **13**, 251–263.

Zajong, R. B. (1965). Social facilitation. *Science*, **149**, 269–274.

(Received Feb. 1, 2002; accepted March 2, 2002)