

# Group size: A possible distorting factor in learning experiments

A. SWERTS, R. PEETERS, and G. d'YDEWALLE  
*University of Leuven/Louvain, Leuven B-3000, Belgium*

Group size may have confounding effects on some dependent variables in experiments conducted in a collective setting. Three experiments revealed unexpected effects of group size on self-paced study and test time, although this was not the main purpose of the studies. Experiments 1 and 2 suggest that the study time of female subjects increases with the group size up to five subjects and then decreases as the group size increases to nine subjects (as used in Experiment 2). When we considered the first subject to finish studying within the group, we noted the same relationship. The phenomena were not found with the male subjects. The within-groups variability of the study time was much larger for the female than for the male subjects. Some of the findings from Experiments 1 and 2 were confirmed in Experiment 3, notwithstanding the much smaller number of subjects and a very different experimental setting. The test times were examined and, again, effects of group size were often obtained. We report these findings in order to make experimenters sensitive to possibly confounding effects from group size when carrying out experiments in a collective setting.

Running experiments with a number of subjects together in the same room saves considerable time for the researcher. In contemporary research on learning and memory, it is implicitly accepted that the data are not affected by the collective setting. Because of the convenience of computers, it has become common practice to run several subjects together, the actual number depending on how many show up for the session. However, a survey of recent literature revealed a virtual absence of publications in which the effect of group size as a potentially noisy variable on the dependent measures is taken into account. In view of the older studies on audience effects in social learning (for review, see Geen & Gange, 1977), this lack of methodological care and completeness in reporting is rather surprising.

The present paper describes three experiments that were not designed for investigating the effect of group size, but in which group size did emerge as a significant factor. In each experiment, it was expected that the main variable (either test expectancy in Experiments 1 and 2 or the experimenter's characteristics in Experiment 3) was not powerful enough to reveal major effects with a small number of observations. In addition, it was not possible to have many observations with the same subjects. Therefore, we decided to collect data from a fairly large number of subjects. To save time, a number of subjects (always from the same sex) were invited at the same time. In all the experiments, the study time was self-paced. Since there were several groups of subjects in each condition, we had to decide to

use either the between-groups or the within-groups variance for the appropriate error term in the ANOVA. In Experiment 1, it became immediately clear that the between-groups variability was larger than the within-groups variability for the female subjects, and a closer look at the data revealed quite systematic effects of group size with female subjects.

## EXPERIMENTS 1 AND 2

In Experiments 1 and 2, the relationship among expectations about the type of test (either open questions or multiple-choice test), the amount of time used to learn a text, and the subsequent performance on open questions and on multiple-choice questions was investigated. The main results and their interpretation (without reporting the effects of group size) are described in d'Ydewalle, Swerts, and De Corte (1983). As the procedure was almost identical for both of these experiments, we will present them together.

## Method

**Subjects.** In Experiment 1, 240 1st-year students from the School of Law at the University of Leuven (Belgium) volunteered to participate. For each experimental session, four to five students of the same sex were invited, but as the number that showed up varied, the experiment was actually run with group sizes ranging from one to five. In Experiment 2, about 500 volunteers from several departments of the University of Leuven participated. The group size varied from one to nine students, each group again being composed of students of the same sex.

**Material.** The same material was used for both experiments. It consisted of two history texts of about equal length and of the same level of difficulty. Each text included the names of 10 important persons, 10 historical dates, and 10 major events in the history of the late Roman Empire. For each text, 30 questions were formulated. There were two versions of each question:

Requests for reprints should be sent to Géry d'Ydewalle, Department of Psychology, University of Leuven, Leuven B-3000, Belgium.

an open question and a multiple-choice question. The multiple-choice questions were exactly the same as the open questions, but four response alternatives were given, one of which was correct. Details about the texts and questions can be found in d'Ydewalle and Rosselle (1978).

**Procedure.** In Experiment 1, the subjects were given 4 min to study Text 1. For half of the groups, the text was followed by multiple-choice questions; the remaining groups received open questions. The subjects were led to believe that the same type of questions would follow after Text 2. However, only half of the subjects received the expected type; the others received the unexpected type (either multiple-choice or open questions). The study time for Text 2 was self-paced. Experiment 2 was identical to Experiment 1, with the exception of the study time for Text 2: One-third of the groups again were given 4 min to learn the text, one-third were given 10 min, and no time limit was set for the remaining one-third.

**Results**

The main hypothesis that subjects consider open-question tests more demanding than multiple-choice tests was confirmed. In both of these experiments, the subjects expecting open questions used longer study times and performed better after Text 2 than those expecting multiple-choice questions (see d'Ydewalle et al., 1983).

Examination of the appropriate error terms in the several ANOVAs revealed that the within-groups variability of the female subjects was much larger than that of the male subjects for Experiments 1 and 2 [ $F(88,74) = 1.47, p < .05$ , and  $F(191,202) = 1.62, p < .001$ ].

We then performed a trend analysis on the average study time of the various sizes of groups (in Experiment 1, from one to five, and in Experiment 2, from one to nine). The averages are plotted in Figure 1. In Experiment 1, a strong linear trend component appeared with the female groups [ $F(1,39) = 12.01, p < .01$ ], but no significant effects were found with the male subjects. From Figure 1, it is apparent that the female

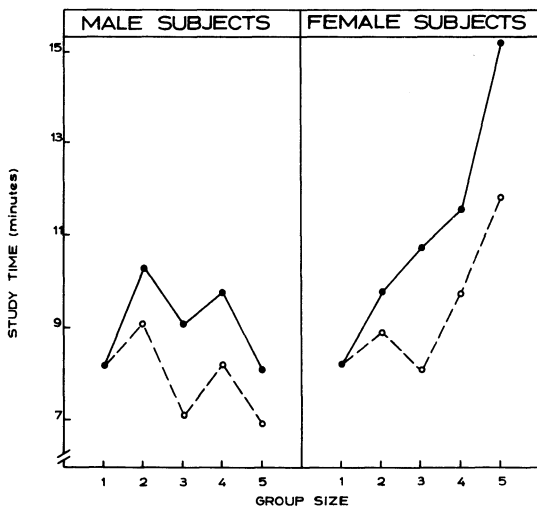


Figure 1. Study time of the groups (solid lines) and of the first subject (broken lines) in Experiment 1.

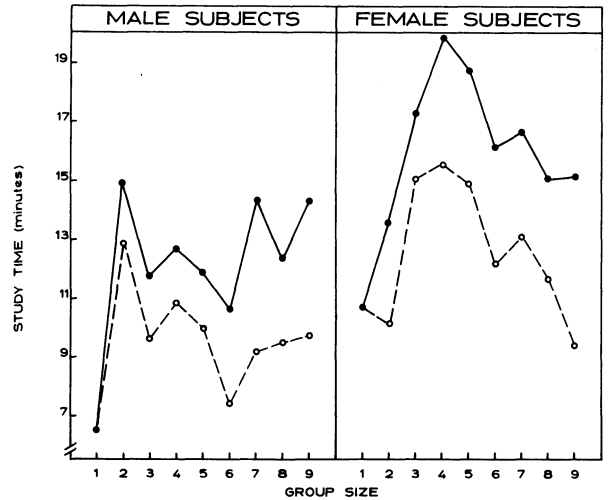


Figure 2. Study time of the groups (solid lines) and of the first subject (broken lines) in Experiment 2.

subjects use more study time when the group size increases. As group size varied from one to nine in Experiment 2, more trend components could be detected. Only a quadratic component with female subjects reached the significance level [ $F(1,39) = 4.36, p < .05$ ]. Inspection of Figure 2 indicates that the linear increase of the study time by the female subjects of Experiment 1 was replicated in Experiment 2. However, a decrease beyond a group size of four subjects was apparent, which produced the quadratic relationship.

We then checked to see if group size affected the study time of the first subject to finish studying in each group. Accordingly, Figures 1 and 2 also give the study time of the first subjects who spontaneously stopped studying Text 2. Figure 1 of Experiment 1 indicates that the linear component with the female subjects again appeared as with the group average, although it was only marginally significant [ $F(1,29) = 3.06, p < .09$ ]. In Experiment 2, the same significant quadratic component as emerged with the group average was significant with the female subjects [ $F(1,39) = 6.35, p < .025$ ] (see Figure 2). No other significant effects were found.

**EXPERIMENT 3**

The study was originally planned to investigate the potential effects of the perceived personality features of the experimenter upon subjects' text learning (for more details on the experiment, see Peeters, 1982). The subjects were led to believe that the experimenter was going to give them a personal examination about the text they had to learn.

**Method**

**Subjects and Material.** In this experiment, 79 1st-year students from various departments at the University of Leuven (Belgium) participated. For each session, five subjects of the

same sex were invited. As not all of them showed up, the actual group sizes ranged from one to four (males) and from one to five (females). Only one group was assigned to each level of the 2 by 2 by 2 by 2 design.

The texts to be learned were modified and abridged Dutch versions of texts by Hess (1965) and Kleitman (1960), both of which were published in *Scientific American*. Their subject was psychology. They were supplied to the subjects in a booklet that included the 15 multiple-choice questions and 1 question concerning the order in which the various topics appeared in the text. The instructions were also printed in the booklet.

**Procedure.** After their arrival in the laboratory room, the subjects were exposed to some "stimulus information" that consisted of a description of their experimenter's personality. These impressions were given in function of varying social and intellectual characteristics (Rosenberg, Nelson, & Vivekananthan, 1968; Zanna & Hamilton, 1972). The induction was made by a conversation between two staff members witnessed by the subjects. The subjects were asked to memorize one of the two texts while anticipating an oral examination by the experimenter shortly thereafter. They were allowed to proceed through the booklet at their own rate. The multiple-choice questions were presented as a means to help the experimenter to prepare more adequate questions for the final oral examination. The study time and the time spent in answering the multiple-choice questions were measured without the subjects' knowledge. The experimenter then fully explained the purpose of the experiment and the necessity of the deception.

## Results

The main hypotheses about the effects of the social and intellectual characteristics of the experimenter upon the study time and test performance could not be confirmed, as no significant effects emerged from the analyses. Again, as a side issue, we looked at the study time and at the time used to answer the multiple-choice questions (test time) as a function of group size. This time, there was no difference between the within-groups variability of the male and female subjects, with  $F$  close to unity. The test time, however, produced considerably less variability for the females than for the males [ $F(18,29)=2.05, p<.05$ ]. As the test time was inversely related to the study time ( $r=-.27$ ), the larger test time variability within the male groups could be considered the inverse of the larger study time variability of female subjects in Experiments 1 and 2.

In the trend analyses on the average study time of each group, a linear trend component emerged with both the male and the female subjects [ $F(1,30)=4.17, p<.05$ , and  $F(1,40)=6.77, p<.025$ ]. A significant quadratic function was also obtained for the male subjects. As can be derived from Figure 3, the linear increase accounts for 44% of the explained trend variance of the male subjects and 67% for the female subjects. Although the pattern of results in Figure 3 does not seem to be as reliable as those in Experiments 1 and 2, the stronger linear increase for the female subjects is in agreement with the data of the first two experiments. A trend analysis on the study time of the subject finishing first in his/her group failed to show any reliable effect.

The relationship between group size and the average test time is also shown in Figure 3. Again, a trend

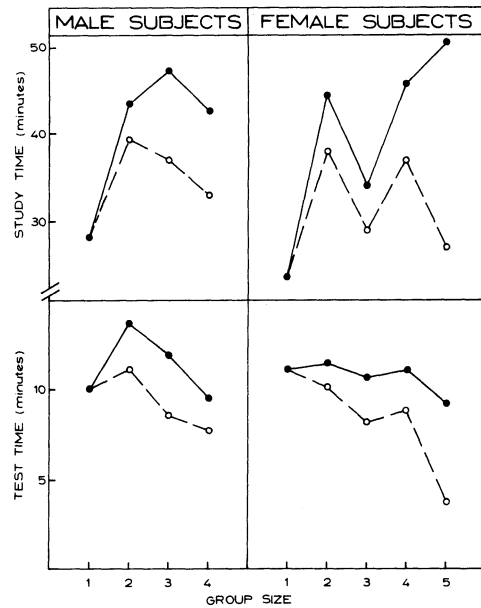


Figure 3. Study and test times of the groups (solid lines) and of the first subject (broken lines) in Experiment 3.

analysis on the average test time did not reveal a significant  $F$  value. The test time of the subject who first finished the entire test showed a significant negative linear and quadratic relationship for the female subjects, mainly due to the extraordinarily short test time of the subject who finished first in the group of five subjects [ $F(1,11)=10.15, p<.01$ , and  $F(1,11)=21.60, p<.001$ , respectively].

## DISCUSSION

We examined the effects of group size on study time because the error variance calculated from the study time of the group means was considerably larger than the error variance estimated from the within-groups variability. As this difference did not appear in the various test performances (either open or multiple-choice questions in Experiments 1 and 2, and multiple-choice questions and topic ordering in Experiment 3), there were no reasons to perform the same group-size analyses for test achievement. Although our data suggest no influence of group size on test achievement, study time, particularly that of the female subjects, is clearly a function of the group size. Moreover, notwithstanding the small number of subjects involved in Experiment 3, some relationships were obtained between group size, on the one hand, and study and test time, on the other.

These phenomena may be only artifacts due to the nature of our group sampling. The size of the groups varied because some subjects did not keep their appointments. The sampling procedure, however, cannot be the only cause. First, the phenomena occurred more strongly with the female subjects. Second, the selection artifact cannot explain in a simple way some observed quadratic relationships. In addition, in Experiments 1 and 2, more within-groups variability on study time was found with the female subjects, and more within-groups variability on test time was found with the male subjects (Experiment 3).

We considered some theories in social psychology for an explanation of our findings. Likely candidates were theories on

social facilitation (Cottrell, 1972; Zajonc, 1965), on bystander effects (Latané & Darley, 1970), and on conformity (Asch, 1952). It became clear that no one theory could explain the overall pattern of our findings. Whatever the underlying processes may be, it is clear that close attention must be given to the effect of group size in learning experiments, especially when the study time is self-paced. Unexpected effects of group size were warned against 20 years ago by Thomas and Fink (1963), but researchers do not seem to have paid much attention to them.

#### REFERENCES

- ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.
- COTTRELL, N. B. Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology*. New York: Holt, Rinehart, & Winston, 1972.
- d'YDEWALLE, G., & ROSSELLE, H. Test expectations in text learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press, 1978.
- d'YDEWALLE, G., SWERTS, A., & DE CORTE, E. Study time and test performance as a function of test expectations. *Contemporary Educational Psychology*, 1983, 8, 55-67.
- GEEN, R. G., & GANGE, J. J. Drive theory of social facilitation: Twelve years of theory and research. *Psychological Bulletin*, 1977, 84, 1267-1288.
- HESS, E. H. Attitude and pupil size. *Scientific American*, 1965, 212, 46-54.
- KLEITMAN, N. Patterns of dreaming. *Scientific American*, 1960, 203, 82-88.
- LATANÉ, B., & DARLEY, J. M. *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts, 1970.
- PEETERS, R. *Leren en memoriseren in functie van de gepercipieerde persoonlijkheid van de proefleider*. Unpublished master's thesis, University of Leuven/Louvain, 1982. (in Dutch)
- ROSENBERG, S., NELSON, C., & VIVEKANANTHAN, P. S. A multi-dimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, 1968, 9, 283-294.
- THOMAS, E. J., & FINK, C. F. Effects of group size. *Psychological Bulletin*, 1963, 60, 371-384.
- ZAJONC, R. B. Social facilitation. *Science*, 1965, 149, 269-274.
- ZANNA, M. P., & HAMILTON, D. L. Attribute dimensions and patterns of trait inferences. *Psychonomic Science*, 1972, 27, 353-354.

(Received for publication January 13, 1983.)