

Effect of Environmental Context on Performance in a University Examination: A Null Result*

Michael C. Corballis

Department of Psychology, University of Auckland

In two large courses in Introductory Psychology, the students who sat their final examinations in the rooms in which they had been lectured scored no higher than those who were assigned to different locations. This fails to support the idea that recall (but not recognition) is susceptible to the effects of environmental context, and adds to other evidence that such effects are not always reliable.

Under some circumstances at least, memory is better when subjects are tested in the location in which they had learned the material than when they are tested in some other location (e.g., Canas & Nelson, 1986; Godden & Baddeley, 1975; Smith, Glenberg, & Bjork, 1978). This phenomenon is an example of the more general influence of context on memory, which is in turn an important aspect of theories of coding and retrieval in memory (Baddeley, 1982; Kintsch, 1974; Metcalfe & Murdock, 1981; Tulving, 1983).

However not all investigators have found that environmental context has an effect, and the results are sometimes in conflict. A common pattern has been that environmental setting affects recall but not recognition, suggesting that contextual cues might serve the purpose of generating the items to be recalled, a process that is not involved in recognition (Godden & Baddeley, 1975; Smith et al., 1978). There are exceptions to this pattern, though. For instance, Canas and Nelson (1986) found that recognition was impaired when subjects were given a surprise test in a location in which they did not expect to be tested. Conversely, Fernandez and Glenberg (1985) failed to find any effect of environmental context on recall in a whole series of experiments, including a virtual replication of an earlier experiment by Smith (1979) that had shown an effect. The effects of environmental context on memory may therefore

be more elusive and idiosyncratic than previously supposed.

The role of the environment in memory may have important practical applications. For instance, Godden and Baddeley (1975) gave underwater divers words to learn either on shore or ten feet under water, and found that recall was better if testing took place in the learning environment. As practical consequences of this result they note, first, that training of divers should not rely too heavily on procedures taught on dry land, and second, that if divers are used as examiners they should record their observations under water rather than rely on later recall on land.

The effects of environment may also have implications for the scheduling of school or university examinations, although the evidence is again conflicting. Some authors have reported that performance is better if students sit tests or examinations in the room in which they had been lectured than if they sit them elsewhere (Abernathy, 1940; Farnsworth, 1934; Metzger, Boschee, Haugen, and Schnobrich, 1979). In a large-scale study involving several big first-year courses, however, Saufley, Otaka, and Bavaresco (1985) found no effect whatever of examination location on examination results.

In 1986, the University of Auckland unwittingly provided a further opportunity to test for the effects of environmental context by scheduling some students in Introductory Psychology to sit their final examinations in the lecture theatre in which they had been lectured, while others were located in other rooms. This article examines the effect of location on the examination results.

*I thank Dianne McCarthy and Keith Macky for making the data available to me. I also thank Murray J. White and one anonymous reviewer for helpful comments. Address for correspondence: M. C. Corballis, Department of Psychology, University of Auckland, Private Bag, Auckland, New Zealand.

Method

The data were obtained from two large undergraduate classes in Introductory Psychology.

Group G

This group comprised students in a lecture course in general psychology, with an enrolment of 631 students (229 men, 402 women). The course involved two one-hour lectures and an optional one-hour tutorial per week for 26 weeks. There were four one-hour tests during the year, and a three-hour final examination. The third of the four tests consisted of multiple-choice items, while the remaining three consisted of questions requiring short written answers. The final examination was made up entirely of written answers and was thus predominantly a measure of recall rather than of recognition.

The class tests were held in the lecture theatre, while the final examination was held in seven different locations scattered across the campus; one of these locations was the theatre in which lectures (but not tutorials) had been held. The students were assigned alphabetically to the different locations, as shown in Table 1.

The marking was carried out by a team of markers, each assigned to one or more questions. However each marker worked through all of the examination scripts, so that there was no confounding of marker with location.

Group E

This group comprised students in a course in experimental psychology, with an enrolment of 403 students, most of whom were also enrolled in the general course described above. This course consisted of one one-hour lecture per week for 26 weeks, and a weekly two-hour laboratory for 20 weeks. On-course grades were based on two one-hour tests held during the year, and on accumulated laboratory marks.

The final examination was held in five different locations, which again included the lecture theatre, and students were again assigned

alphabetically to these locations, as shown in Table 1. This group was less critical than Group G to the hypothesis under study since a high proportion of their activities during the year took place in laboratories rather than in the lecture theatre. Marking was carried out in the same manner as for Group G.

Results

Group G

The mean scores on the examination are shown as a function of examination location in Table 2. Location C was the theatre in which the students were lectured, and the mean score for the group in this location lay between the maximum and minimum scores for the other locations, and was very close to the mean for the class as a whole. Analysis of variance revealed no significant effect due to location ($F(6,627) = 0.90$).

Since the distribution of men and women, and the distribution of students across faculties, fluctuated somewhat between groups, a second analysis of variance was carried out in which gender and faculty were included as (non-orthogonal) factors. The faculties were Arts (301 students), Science (190 students), law (72 students), and Others (68 students); this last category comprised Architecture, Music, Commerce, and Optometry, each with very small enrolment, and a few students not taking full degrees. Each effect in this analysis was adjusted for the remaining effects, so that the effect of location was assessed with gender and faculty, and the interaction between them, as covariates.

Again there was no significant effect of location $F(6,578) = 1.44$. Adjusted means are also shown in Table 2. These means were least-squares estimates of what the means would have been had the groups in each location been matched for faculty and gender. Note that the mean for location C is adjusted slightly downward but is almost exactly the same as the mean for the class as a whole.

Two effects in this second analysis did prove significant, however: Women scored significantly more highly than men, $F(1,578) = 4.03, p < .05$, and there was a significant difference between faculties, $F(3,578) = 3.39, p < .02$. The actual and

Table 1: Allocation of students in each course by surname, or initial letter of surname, to examination locations

Location	Group G	Group E
A	A — Coad	
B	Cochrane — G	
C*	H — L	Forsyth — L
D	M — O	
E	P — Q	
F	R — S	A — Foote
G	T — Z	P — S
H		M — O
I		T — Z

*Lecture location

Table 2: Means, standard deviations, and adjusted means of examination marks (out of 100) for each examination location, for all of Group G

Location	N	M	SD	Adjusted Mean ^a
A	104	55.49	16.3	58.39
B	95	57.61	16.1	59.32
C ^b	110	55.97	16.7	55.17
D	90	55.54	14.6	57.18
E	47	52.61	17.8	57.04
F	93	55.63	17.3	57.34
G	92	52.87	17.3	48.46
Overall mean 55.19				

^aAdjusted for variations in faculty and gender.

^bLecture location.

adjusted means are shown in Table 3. There were no significant interactions.

In the next analysis, the four test scores were included as covariates. Since not all students sat all four tests, the sample was reduced to 524 students. These students were probably more likely than the class as a whole to have attended the lectures, and might therefore be expected to provide a stronger test of the effects of environmental context. Again, however, there was no significant effect of location, $F(6,464) = 1.76$. Means and adjusted means are shown in Table 4; note that the values for location C are again between the maxima and minima for the different locations, and are very close to the overall mean.

This analysis also failed to reveal any significant effects of gender or faculty, or any significant interactions. However all four class tests proved to be significant predictors of the examination mark, $F(1,464) = 47.04, 46.02, 13.28,$ and 64.58 , respectively, $p < .003$. The correlations of these tests with the final examination mark were .62, .65, .41, and .62, respectively. It is of some interest that the third of these tests, the multiple-choice test, was the poorest predictor, perhaps because there were no multiple-choice items on the final examination.

Group E

The mean scores on the final examination for the five locations are shown in Table 5, and again the effect of location was not significant, $F(4,397) = 0.47$. The mean for location C, which was where the students were lectured, was again very close to the overall mean. In a second analysis, gender and faculty (Arts, Science, Law, and

Table 3: Means and adjusted means of examination marks for each gender and faculty in Group G.

Category	N	M	Adjusted Mean
Men	229	52.33	54.46 ^a
Women	402	57.03	57.80 ^a
Arts	301	54.43	52.16 ^b
Science	190	53.73	53.15 ^b
Law	72	57.43	58.44 ^b
Other	68	61.47	60.76 ^b

^aAdjusted for variations in location and faculty.

^bAdjusted for variations in location and gender.

Table 4: Means and adjusted means of examination marks (out of 100) for each location for those students in Group G who completed all on-course tests

Location	N	M	Adjusted mean ^a
A	83	58.46	58.24
B	82	58.79	58.64
C ^b	96	56.42	56.69
D	73	58.08	58.14
E	38	55.68	56.35
F	75	60.54	59.14
G	78	54.03	55.06
Overall mean 57.52			

^aAdjusted for variations in gender, faculty, and test scores.

^bLecture location.

Table 5: Means and adjusted means for examination marks (out of 100) for each location for Group E

Location	N	M	Adjusted mean ^a
F	100	53.76	55.36
C ^b	103	53.20	51.89
G	50	48.74	50.93
H	94	53.35	57.67
I	55	53.24	48.76

^aAdjusted for variations in gender, faculty, and laboratory mark.

^bLecture location.

Others) were included as factors, and the laboratory mark was included as a covariate. The adjusted means are shown in Table 5. This analysis did produce a significant effect of location ($F(4,361) = 2.81, p < .05$), but this seems to be due to the exceptionally high adjusted score for those in location H and the exceptionally low score for those in location I. The allocation of students to locations was of course nonrandom, but was determined by alphabetic ordering of surnames, and it is conceivable that this introduced a bias. Those whose surnames began with the letters T through Z were assigned to location I, and informal inspection suggested that a disproportionately high number of these names were of Asian or Polynesian origin. It is possible that immigrant or minority students were disadvantaged.

Pairwise *t*-tests indicated that the adjusted mean for location H differed from those of locations C, G, and I, and that the adjusted mean for location I also differed from that of location F ($p < .05$); none of these comparisons is significant, however, if a Bonferroni correction is applied (i.e., if the *p*-values are multiplied by 10). These differences in any case do not seem to bear on possible context effects associated with location C.

The effect of gender was not significant, $F(1,401) = 0.98$, but there was a significant effect of faculty, $F(3,401) = 2.64, p < .05$. Means and adjusted means are shown in Table 6. The laboratory mark proved to be a strong predictor of the examination mark, $F(1,401) = 400.76, p < .0001$.

Discussion

The results fail to show any evidence for an effect of environmental context on marks in an examination that depended primarily on recall of information. The locations themselves varied quite widely in their physical characteristics, so that it was not unreasonable to have expected features specific to the lecture location to have become associated with the course content and to have assisted retrieval. The lecturers themselves, who might also have provided context cues, were not present during the examination. However Abernathy (1940)

Table 6: Means and adjusted means for each faculty, in Group E

Faculty	<i>N</i>	<i>M</i>	Adjusted mean ^a
Arts	210	49.88	50.24
Science	114	55.25	54.47
Law	33	52.18	53.16
Other	45	60.87	53.73

^aAdjusted for variations in location, gender, and laboratory mark.

found that the presence or absence of the course instructor had less of an effect on examination scores than did environmental context, although both factors did have an influence.

The present results essentially confirm those of Saufley et al. (1985), who similarly failed to find any effect of environmental context on examination scores in several introductory courses with large enrolments. However they contrast with those of Metzger et al. (1979) and of Abernathy (1940), where context effects did occur. These last two studies involved smaller samples, but otherwise more powerful designs. Metzger et al. (1979) used a crossover design with two streams of 25 from a class of 50. The students were given class tests in their usual classroom on three consecutive weeks, as a baseline; on the fourth week one stream was tested in a different room while the second stream remained in the usual room, and on the fifth week the first stream remained behind while the second shifted to the other room. Changing rooms had a significant detrimental effect on test scores. Abernathy (1940) used a complex Latin Square design in which fourth- and fifth-year classes were divided into four groups, matched for their performance in earlier years, and were then tested four times with counterbalanced conditions involving same or different room and the instructor present or absent. Performance was adversely affected both by the change from the usual room, and by the absence of the usual instructor.

In these last two studies, too, testing was more directly a part of on-going teaching. In the present study, by contrast, the final examination covered material going back over the whole academic year, and there was a gap of several weeks between lectures

and the examination, with the most intensive period of study probably taking place during those weeks and outside of the lecture theatre. Similar conditions no doubt prevailed in the study by Saufley et al. It is possible, then, that the absence of context effects was due not so much to experimental design as to differences in the conditions of learning and testing.

The precise conditions under which environmental context plays a role in recall nevertheless remain unclear (cf. Fernandez & Glenberg, 1985). For the present, however, there is little reason to believe that any bias is introduced if some students are scheduled to sit final examinations in the lecture room, while others are scheduled elsewhere, at least in large introductory courses in which the bulk of the material is also contained in a textbook.

References

- Abernathy, E. M. (1940). The effect of changed environment conditions upon the results of a college examination. *Journal of Psychology*, *10*, 293-301.
- Baddeley, A. D. (1982). Domains of recollection. *Psychological Review*, *89*, 708-729.
- Canas, J. J., & Nelson, D. L. (1986). Recognition and environmental context: The effect of testing by phone. *Bulletin of the Psychonomic Society*, *24*, 407-409.
- Farnsworth, P. R. (1934). Examinations in familiar and unfamiliar surroundings. *Journal of Social Psychology*, *5*, 128-129.
- Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, *13*, 333-345.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, *66*, 325-331.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, N. J.: Erlbaum.
- Metzger, R. L., Boschee, P. F., Haugen, T., & Schnobrich, B. E. (1979). The classroom as a learning context: Changing room affects performance. *Journal of Educational Psychology*, *71*, 440-442.
- Metcalfe, J., & Murdock, B. B., Jr. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning & Verbal Behaviour*, *20*, 161-189.
- Saufley, W. H., Jr., Otaka, S. R., & Bavaresco, J. L. (1985). Context effects: Classroom tests and context independence. *Memory & Cognition*, *13*, 522-528.
- Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning & Memory*, *5*, 460-471.
- Smith, S. M., Glenberg, A. M., & Bjork, R. A. (1978). Environmental context and memory. *Memory & Cognition*, *6*, 342-353.
- Tulving, E. (1983). *Elements of episodic memory*. London: Clarendon Press/Oxford University Press.