



Distributed practice learning foreign
vocabulary – The effects of interval
length

*Practice with longer intervals means
less work and better learning for the
long term*

Summary report of an experimental study

Catherine O. Fritz, Rachel Duffy and Peter E. Morris
May 2010

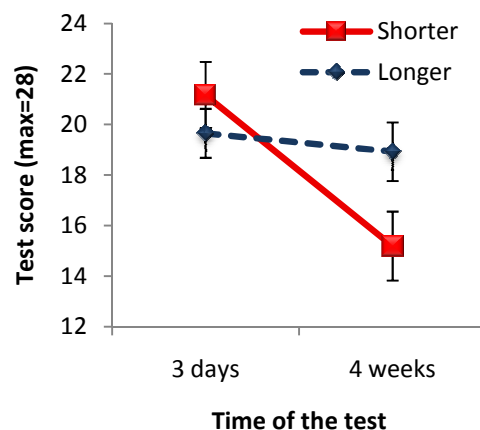
Practice with longer intervals means less work and better learning for the long term

By spacing out practice with longer gaps, students can work less and learn better. University students using phase-6®, a web-based system for helping learners to practice what they're learning, either practised with shorter (1, 2, 3, and 4 days) or longer (2, 5, 9, and 15 days) intervals while learning 28 Finnish to English word translations over 10 days; there were 26 students assigned to each condition. Longer practice intervals led to better learning as measured by less forgetting and performance on a test 4 weeks later. People reported similar positive experiences with phase 6 regardless of the length of the intervals.

Participants who studied with longer intervals did less work. They had less practice¹ than those with shorter intervals. The measure of practice is the number of times that they were asked to translate a word over the 10 day period.

Four weeks later, the people who studied with longer intervals knew more. All participants were tested twice following the 10 day practice period. The first test occurred 3 days later and the second a further 4 weeks later. Test performance is summarized in Figure 1. Overall, some forgetting occurred between the two tests, but people with longer practice intervals forgot little, if anything, whereas those with shorter practice intervals forgot a substantial number of words².

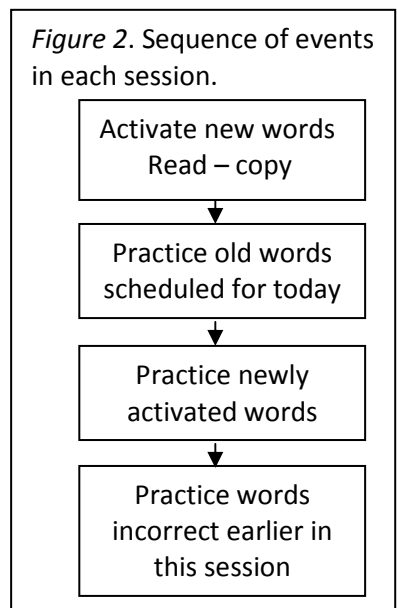
Figure 1. Translation test scores 3 days and 4 weeks following 10 days of practice. The test provided all 28 Finnish words; the score for each participant was the number of correct English translations provided.



¹ With longer intervals, $M = 180$, $SD = 50.2$, $N = 26$; with shorter intervals, $M = 212$, $SD = 45.8$, $N = 26$. The difference was moderately large, Cohen's $d = .7$, and was statistically significant, $t(50) = 2.45$, $p = .02$, two tailed.

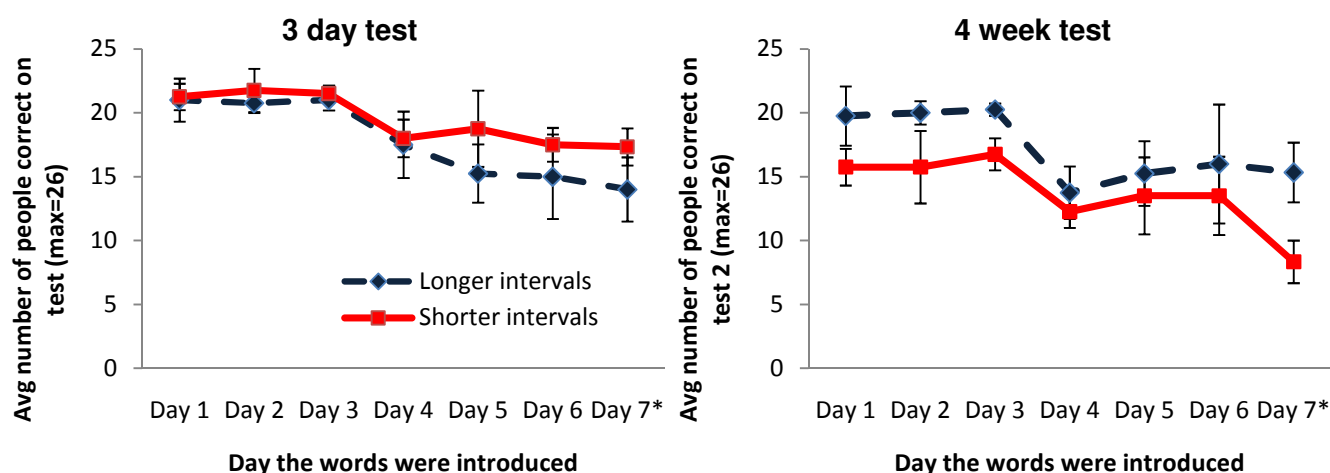
² Overall, scores on the later test were significantly lower than those on the earlier test, $F(1,50) = 17.52$, $MSE = 16.62$, $p < .001$, $\eta_p^2 = .26$, but there was no overall effect of interval length, $F(1,50) = 0.54$, $MSE = 60.16$, $p = .47$, $\eta_p^2 = .01$. The interaction between interval length and the time of the test was significant, $F(1,50) = 0.70$, $MSE = 16.62$, $p = .002$, $\eta_p^2 = .18$. For the earlier test, shown on the left side of Figure 1, the two groups did not differ significantly, $t(50) = 0.91$, $p = .37$, $d = .2$, whereas scores on the later test, on the right side of Figure 1, were substantially and significantly higher following practice with longer intervals, $t(50) = 2.09$, $p = .04$, $d = .6$. Another way to describe the interaction is that with shorter practice intervals - the dark lines in Figure 1 - there was considerable loss during the 4 weeks between the tests, $t(25) = 5.78$, $p < .001$, $d = 1.1$, whereas with longer practice intervals - the pale lines - the loss was negligible, $t(25) = 0.60$, $p = .56$, $d = .1$.

The study sessions introduced four words each day for seven days with three further days for practice without new material. To activate a word, the translation appeared and the participant copied the English word. Following activation of new words, previously practised words that were scheduled for additional practice appeared. The Finnish word appeared and asked for the English translation. If the answer was correct, the word moved up to the next phase – that is, the interval until the next practice session for that word was extended. If the answer was incorrect or missing, the correct answer appeared, the participant types the English translation, and the word appeared again during the same session (phase 1). Following the practice on previously practised words, the newly activated words were practiced, along with words that had been translated incorrectly earlier in that session. The session ended when all of the words scheduled for that day had been translated correctly.



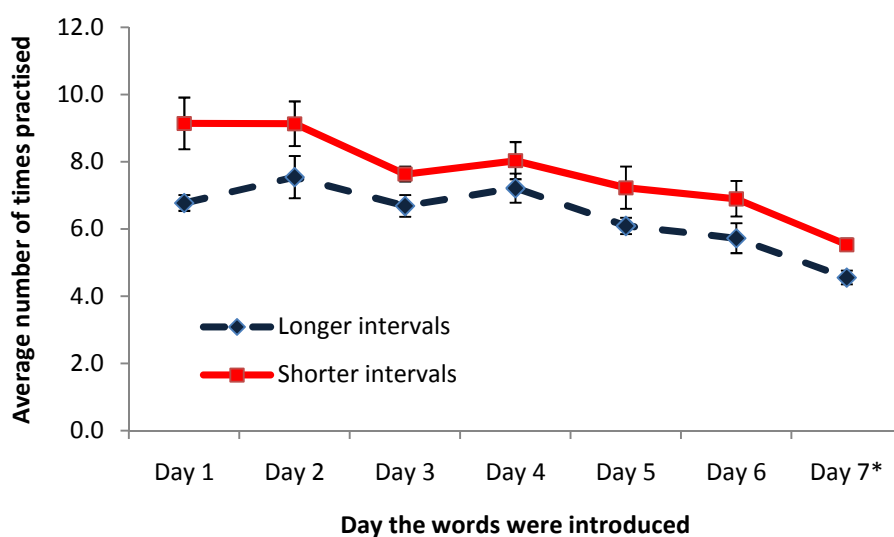
Of the 28 words, one was a cognate (*siemen* for *seed*) which was more easily learned by participants. That word was introduced on the 7th day, and has been excluded from the following summaries. A brief description of the test performance based on the day the word was introduced is provided in Figure 3. Notice that the words introduced on the first three days were somewhat more likely to be correctly translated than the words introduced later; perhaps future research should investigate the possible benefits of introducing new material for a few days at a time with a few intervening days of practice without new material. Interval length did not significantly affect test performance on the 3-day test regardless of the day that the words were introduced; on the 4-week test, the people who practiced with longer intervals were more likely than the people who practiced with shorter intervals to recall the words introduced on day 1, 2, 3, and 7.

Figure 3. Test performance based on the day words were introduced. Each point (except Day 7) represents the number of people who translated the word correctly, averaged across the four words introduced on that day. For Day 7, only three words were included; performance on the cognate was artificially high so it was removed. Error bars represent +/- 1 SEM.



One possible explanation for a benefit to words introduced earlier is that they have more opportunities for practice. Figure 4 shows the average number of practice events for words based on the day they were introduced, reported separately for the two groups. People who practiced with shorter intervals clearly had more opportunities for practice and words introduced earlier were more often practised.³

Figure 4. Mean frequency of practice for the words introduced each day. The frequency of practice was calculated for each word, as an average across people, and then the four words for each day (except Day 7 with three) were averaged to produce these data points. The error bars represent +/- 1 SEM across the four words.



Participants' impressions.

Following the first test, participants were asked to rate their experience with the phase-6[®] learning experience. Table 1 summarizes their ratings of difficulty; two people did not return the questionnaires. Generally people reported that the learning process was moderately difficult throughout the investigation. It is noteworthy that the two groups did not differ significantly in their post hoc perceptions of difficulty⁴ despite the obvious difference in absolute difficulty as evidenced by the need for more practice and the presence of more errors during practice for the group with

³ The observed difference between the two groups was significant, $F(1,50) = 7.38$, $MSE = 21.05$, $p = .009$, $\eta_p^2 = .13$, as was the effect of day, $F(6,300) = 135.35$, $MSE = 1.84$, $p < .001$, $\eta_p^2 = .41$; the interaction was marginally non-significant, $F(6,300) = 1.93$, $MSE = 1.84$, $p = .08$, $\eta_p^2 = .04$. For both groups, a linear trend was clearly observed with respect to days, overall $F(1,50) = 109.34$, $MSE = 2.94$, $p < .001$, $\eta_p^2 = .69$; for practice with shorter intervals, $F(1,25) = 71.15$, $MSE = 3.19$, $p < .001$, $\eta_p^2 = .74$; for practice with longer intervals, $F(1,25) = 39.37$, $MSE = 2.69$, $p < .001$, $\eta_p^2 = .61$. A quadratic trend was also significant overall, $F(1,50) = 9.39$, $MSE = 2.94$, $p = .004$, $\eta_p^2 = .16$, and for practice with longer intervals, $F(1,25) = 7.85$, $MSE = 2.69$, $p = .01$, $\eta_p^2 = .24$, but not for practice with shorter intervals, $F(1,25) = 1.77$, $MSE = 3.19$, $p = .196$, $\eta_p^2 = .07$.

⁴ $F(1,48) = .436$, $MSE = 4.57$, $p = .44$, $\eta_p^2 = .01$

the shorter intervals. People perceived the learning task as becoming more difficult across the 10 days of the investigation⁵.

Table 1
Post hoc Ratings of Difficulty of the Learning Process

Question (1 = easy, 10 = difficult)	Longer (n = 24)		Shorter (n = 26)	
	Mean	SD	Mean	SD
How hard did you find the learning process <u>at the beginning</u> of the investigation?	4.0	2.40	4.4	2.40
How hard did you find the learning process <u>in the middle</u> of the investigation?	5.6	1.61	6.3	1.71
How hard did you find the learning process <u>at the end</u> of the investigation?	6.3	2.58	6.0	2.31

Regardless of the length of the intervals during study, people reported that they enjoyed using phase-6[®] and that they thought it was a good way to learn foreign language vocabulary; responses are reported in Table 2. Those who thought that phase-6[®] was a good way to learn translations gave reasons such as “lots of practice and repetition” (7 longer + 7 shorter people), “retyping incorrect answers” (1 + 5 people), easy and/or quick (3 + 3 people), personal testing (2 + 3 people). Those who did not think so reported that they could not see how sentences could be learned (1 + 1), stated that their preferred way to learn was by listening (1 + 1), observed that they could cheat or just focus on part of the Finnish word (1 + 1), suggested that they had not learned the words, merely recognized them (2 + 0), stated that it was difficult or frustrating (2 + 1), or suggested that additional cues should be provided to assist learning (1).

Table 2
Post hoc Responses to Questions About the Experience of Using Phase-6

	Longer (n = 24)		Shorter (n = 26)	
	Yes	No	Yes	No
Did you enjoy your experience with phase-6 [®] ?	23	1	24	2
Do you think that phase-6 [®] was a good way to learn translations?	16	8	20	6
Are there any techniques that you can identify that you think helped with your practice sessions – for example, word associations or similarities between the Finnish and English translations?	22	1	23	3

⁵ The main effect of time in the investigation was statistically significant, $F(2,96) = 11.50$, $MSE = 4.96$, $p < .001$, $\eta_p^2 = .19$, as were the linear and quadratic trends. The interaction term was not statistically significant, $F(2,96) = .64$, $MSE = 4.96$, $p = .53$, $\eta_p^2 = .01$.

To help them learn the words, almost everyone reported using word associations or similarities between the words to supplement the effects of practice testing; responses are tallied in Table 2.

Dissemination and impact.

We intend to report this research to a national cognitive psychology conference (British Psychological Society, Cognitive Section Conference) and an international memory conference (Conference of the Psychonomic Society) in 2010 and if support is found for the followup study, will be reported in a major memory journal. We also plan to identify a journal and/or website with a target audience of foreign language teachers, and prepare a summary for publication there.

Future work.

This design of this research project was cut back due to a fixed calendar deadline for reporting the work combined with delays in having access to the programme with the two groups clearly defined and working. Nevertheless, it has provided us with some excellent evidence, beyond that available in the memory literature at present, about the long term value of longer intervals.

The next study we propose to do will involve participants for a longer period of time – probably a month – in order to assure that words are promoted to phase 5 and 6. It will also test participants early, after a month, and again after 6-12 months to look at longer term retention. We will introduce more words – perhaps 40 – and include some useful phrases as well. Rather than comparing shorter to longer intervals, we intend to investigate the effects of uniform vs. expanding intervals (i.e., expanding intervals of 1-2-5-9-14 days compared to uniform intervals of 6-6-6-6-6 days) to address a question that has arisen in the memory literature. We will set up sessions to include practice translating in both directions as well. We will also include a condition with LinkWord-style associations in the early stages to further assist learning if we can recruit a large enough number of participants and identify a way to accomplish this using phase-6. This condition should greatly increase the effectiveness of phase-6®, but may require software support from phase-6®. We hope to run a further study in 2010.

Another question that bears research is one that we have discussed in the past: When an error occurs, should the item drop back to phase 1, or should it drop back by one phase after being correctly practised in the session where the error occurred? The online version of phase 6 that we used takes the former action, but the memory literature suggests that the latter course would be more effective, so it would be worthwhile to investigate this question. More a more recent desktop version allows the option. We hope that future research can access a version of phase-6 that supports either action so that their effects can be compared. These results would help to better inform and develop our understanding of memory and could help to improve phase-6® and the advice that accompanies it.

Longer term plans include expanding the demonstrated effectiveness of phase-6® by using it to support learning in a content area (e.g., cognitive psychology) and in a specialized area (e.g., statistics).