教育研究學報 1995，第十卷第一期，94-98

Educational Research Journal 1995, Vol. 10, No. 1, pp. 94-98

Dubious Benefits of Exposure to a British Environment on Hong Kong Teachers' English Ability: Comment on Dolan (1994)

Kit-tai Hau

The Chinese University of Hong Kong

Dolan (1994) compared the English ability of three groups of Hong Kong (HK) teachers [and a group of British (UK) students] who were attending university courses in UK. Among other things, he concluded that there was general improvement in language proficiency with British exposure. In this article, Dolan's analyses were re-examined. It was found that UK exposure might lead to an increase in HK teachers' specific UK knowledge. However, contrary to Dolan's conclusion, reanalyses showed that HK teachers' general English proficiency had not improved even after nine months of exposure. The benefits of UK exposure were unsubstantiated and dubious.

Dolan (1994) 比較三組在英國大學就讀的香港教師（及一組英國學生）之英語能力，在結論中，他指出這些受試者的一般英語能力會隨他們在英國生活的體驗而有所改善。本文重新檢視 Dolan 的論據，發覺英國的生活可能增加香港教師對英國的認識，但與 Dolan 的結論不同，重新分析顯示即使長達九個月之英國生活，亦不能改善香港教師之英語能力。「英國生活體驗能改善英語能力」之說，在該研究內並無確實數據支持。

In a study which compared three groups of Hong Kong (HK) teachers (and a group of UK students) who had different degrees of exposure to the British environment, among other things, Dolan (1994) concluded, "It revealed that the Hong Kong teachers' recall performance improved as they became more used to the English background and had had time to acquire 'British' schemata" and "it also highlighted the benefits of prolonged periods of exposure to English and general improvements in language proficiency" (p.35) (the general improvement in language proficiency is the main concern of this paper, which is abbreviated as *'Dolan's claim'* below). In the present reanalyses of Dolan's study, it is found that at best, exposure to the British environment may increase one's knowledge of the country and hence helps the recall of British theme items. However, *Dolan's claim* of general improvement in English proficiency through UK exposure is totally unsubstantiated. In this article, other minor methodological problems with Dolan's study are also discussed.

Correspondence concerning this article should be addressed to Kit-tai Hau, Department of Educational Psychology, Faculty of Eduction, the Chinese University of Hong Kong, Shatin, NT, Hong Kong.

As argued by Dolan (1994, see also Carroll, 1986), the English proficiency of HK teachers or their students will be enhanced if they have the chance to stay in HK where the language is spoken as a native language. In view of the recent public concern in HK on how to improve teachers' and students' English ability, *Dolan's claim* has significant policy implications on whether the HK government should spend money on providing chances of exposure (such as supporting HK teachers for short stays in UK). Because of such importance, this article is written to examine *his claim.*

I believe no one will question the benefit of immersing oneself in a culture where English is spoken as a native language. However, two main pragmatic concerns are: 'How long should the immersion be to be effective?' and 'Which aspects of language proficiency are improved most significantly in the immersion?' Obviously, a single day immersion will not help, and a ten year immersion program is not viable. So, an important task is to find the optimum exposure that will substantially enhance language proficiency. Equally critical in the evaluation of the immersion program is that it should improve HK teachers' general language proficiency (in contrast to increasing UK knowledge alone). I am not arguing that specific UK knowledge (e.g., 'The Atlantic often batters the Scottish coast in the winter months

from November to March', 'The M25 motorway was intended to relieve traffic in London but more bottlenecks have resulted' as found in the items tested) is not important. Rather, as this not intended to be a cultural exchange program, the main criterion of success is the improvement in teachers' general English proficiency.

The primary concern of the present paper is to evaluate whether Dolan's data and analyses supported *his claim.* I make no attempt and it is actually quite beyond the scope of this short comment to review and discuss the general advantages and limitations of immersion programs. Nonetheless, it appears that Dolan's study was not primarily concerned with students' general language proficiency. Hence, the conclusion in both his abstract and discussion section about the benefits of British experience on language proficiency can be very misleading and dangerous, especially when it has important educational policy implications.

Naturally I do not assume and expect that each of Dolan's analyses is directly related to general English proficiency. But as Dolan did not explicitly state how he arrived at *his claim*, I have to examine all the analyses presented in his article. The purpose is to see whether any of these analyses individually or when combined can support *his claim.* In brief, my main criticism is that none of Dolan's analyses can support *his claim* about general improvement in language proficiency. Of course, I am not arguing that his analyses do not support any of his other conclusions (e.g., improvement in UK knowledge). Nor am I suggesting all the conclusions in Dolan's article are wrong. Wherever appropriate, I will also make suggestions on how Dolan's analyses can be extended to examine *his claim* of improvement in general English proficiency.

Dolan's Study

Dolan examined subjects' performance on an English recall task. These subjects were provided with 16 sentences — 8 with British (UK items) and 8 with HK background themes (HK items) (e.g., UK item: 'Most families listen to or watch the Queen's address to the nation after Christmas dinner at three o'clock'; HK item: 'Hong Kong young people are often stopped in the New Territories to have their I.D. cards checked'). They listened to audio tapes of these sentences and wrote down what they thought they had heard. Accuracy in spelling was not assessed in the task. His subjects were 64, 40 and 40 HK English teachers who had been living in England for 0, 3, and 9 months (HK_{0m}, HK_{3m}, HK_{9m} groups) respectively and were all taking a B.Ed. course for teachers at Nottingham university. Another group of 20 native UK undergraduates were also assessed with the same task.

Dolan found: (1) the F-values of three separate oneway ANOVAs of the UK, HK, and total (UK+HK) scores by the four groups of subjects were all significant (see Table 1); (2) linear trend comparison of the three HK groups was significant with the UK items but not with the HK ones; (3) regression analyses using HK score to predict UK score, followed by oneway ANOVA by the four groups of the standardised difference score (ipsative score = actual score - predicted score) (for details see Dolan, 1994, p.34) showed that the HK groups with little UK exposure were much worse than predicted, whereas those with longer exposure were better than anticipated.

Table 1
Recall Performance by Degree of British Exposure

Group/ Degree of Exposure	N	Uk themes Mean	Uk themes SD	HK themes Mean	HK themes SD	Difference Score Mean	Difference Score SD
HK (0 month)	64	15.89	4.63	33.62	5.71	-6.20	3.92
HK (3 month)	40	18.22	7.36	29.97	9.11	-1.31	3.48
HK (9 month)	40	25.22	6.72	34.22	8.59	2.72	3.98
native UK	20	39.90	5.23	34.75	6.02	17.02	3.48
oneway ANOVA F-ratio		88.40**		3.01*		195.76**	

Note. Table adapted fromDolan (1994).
*p< .05, **p<.01.

Comment

As relevant to the following comments, it is suffice to argue that subjects' recall performance is determined by their general English proficiency and knowledge of the respective (UK/HK) environment (abbreviated as 'ability' and 'knowledge' respectively in the following). Thus, for example, native UK students did better on the UK items than the HK groups because they had better ability and UK knowledge. In effect, Dolan hypothesised that UK exposure would increase HK teachers' ability and/or UK knowledge. However, as perceived by many HK educational policy makers, the former is more important, that is, their main concern is whether exposure will enhance ability.

In the following, Dolan's (1994) analyses will be critically examined with respect to the information available in his report. Furthermore, the above mentioned criteria will also be used to evaluate whether there is any support for *his claim* of general improvement in English proficiency through UK exposure.

Confounding Effects

In Dolan's two separate ANOVAs of UK and HK scores, the overall F-values were both significant. However, for the first ANOVA, the four groups might differ by their ability as well as their UK knowledge. For the second ANOVA, the UK group differed from the other three HK groups by their HK knowledge, whereas all four groups might also differ by their ability. So, in Dolan's first ANOVA, there was no way to distinguish whether ability or UK knowledge was contributing to the differences. Similarly, in Dolan's second ANOVA, the effects due to HK knowledge and ability cannot be distinguished. Thus, this set of ANOVAs cannot be used to support *Dolan's claim* (i.e., improvement in ability/general language proficiency).

A more relevant analysis, if Dolan were to justify *his claim*, would be an ANOVA of the HK score by the three HK groups. As all the HK groups have similar HK knowledge, the proposed ANOVA will show the difference, if any, due to ability. It will tell us whether ability changes with UK exposure.

Lack of Pairwise Comparsion

Dolan's reliance on the overall F-value alone without further post-hoc multiple comparisons is problematic. Some people (as exemplified by one of the anonymous reviewers) may argue that pairwise comparisons are not necessary because they add nothing to the pattern. This is an invalid argument because multiple comparisons are not used to reverse the direction of the differences. Instead, they help us to interpret the overall significant F-value correctly. A closer examination will show that a large difference between the UK and the HK groups alone (without great differences among the HK groups) can give significant F-values. Thus, post-hoc pairwise comparisons of the means are useful and necessary to locate which pairs of groups are actually different.

The multiple comparison suggested above is particularly important when a visual inspection shows great differences between the UK and HK groups and relative small changes among the HK groups. For the UK items, using information reported by Dolan, it was found that the greatest differences were actually between the UK and HK groups, effect size (ES), UK vs. $HK_{0m, 3m, 9m} = 5.03, 3.22, 2.34$ respectively; whereas those among the HK groups were relatively smaller .40, 1.69, .99 only. Furthermore, as pointed out earlier, due to the confounding effects, there is no way to know whether the differences among the HK groups should be attributed to their different abilities or to their different UK knowledge. Therefore, this analysis cannot be used to support Dolan's claim.

For the HK items, the pairwise comparisons did not show strong consistent patterns. The UK group was not different from the HK_{9m} group (ES=.07), but was better than the HK_{0m} (ES=.20) and HK_{3m} (ES=.58) groups. Once again, it is not certain whether such difference should be attributed to their different abilities or HK knowledge. The differences among the HK groups were inconsistent. Though the HK_{9m} group was better than HK_{3m}(ES=.48), contrary to Dolan's exposure theory, the HK_{0m} group was better than HK_{3m}(ES=.50) and was not much different from HK_{9m} (ES=.09).

The results of the above comparisons on the HK items are consistent with Dolan's linear trend analysis on the HK items with the three HK groups. Both analyses show that there is no clear trend of improvement in general ability with UK exposure. These analyses are most related to *Dolan's claim*, but surprisingly the results are in an opposite direction to what Dolan has concluded. Using the limited available information and the restricted sample, it can only be concluded that HK students' ability does not improve with UK exposure, even after a nine month stay. *Dolan's claim* is completely contradictory to what the data and analyses are showing.

It is worth noting that the above nonsignificant difference in ability with the HK groups when taken together with the result of the first ANOVA suggests that the differences in HK items are mainly due to differences in UK knowledge rather than ability. Once again this particular analysis and result cannot be used to support *Dolan's claim.*

Uninteresting Combined Scale

In view of the disparate trends of the UK and HK items, Dolan's third ANOVA on the total (UK+HK) scale is uninteresting because the result only reflects the massive effect in the first ANOVA (i.e., the UK items). In this particular case, when independent analyses have been conducted on the separate UK and HK scales and provide more mean-

ingful interpretation, the analysis on the combined scale does not give additional useful information. Furthermore, as discussed previously, due to the confounding effects of ability, HK knowledge, and UK knowledge, nothing conclusive can be obtained in this ANOVA as regards the effects of UK exposure on students' ability.

Inappropriate Four Group Comparison

Dolan, in an attempt to 'take account of sampling differences between the Hong Kong groups' (p.34), used the HK score to predict the UK score, followed by an ANOVA on the standardised ipsative score (actual-predicted) (see p.34 for details). As discussed above, the HK score is affected by both the ability and HK knowledge factors. The effect of ability is likely to be positively related to the extent of exposure and is different for all four groups, whereas the HK knowledge factor is constant for all HK samples but has a negative effect on native UK students (i.e., UK students have relatively little HK knowledge). Sampling problems, if any, are confounded with the extent of exposure and cannot be totally isolated in this particular design. Due to these complicated effects on the four groups, therefore, not much can be learnt with respect to students' general ability even after partialling out the effects due to differences in HK scores. Thus, this set of analyses cannot be used to support *Dolan's claim.*

Actually, even for Dolan's originally stated purpose, the main hypothesis of this study (i.e., effects of schematic background/knowledge on recall), the comparison of the three HK groups seems to be more appropriate. The analyses will be identical to those used by Dolan, but three HK groups instead of 3 HK + 1 UK groups (as in his study) will be used (see Table 1 last column). The suggested analyses on the three HK groups will give a clearer comparison and a more definite conclusion of the effects of exposure on UK knowledge. As all these three groups have the same HK knowledge but are of different ability levels (due to different lengths of exposure), partialling out the effect of HK score in the ANOVA analyses of the difference/ipsative score is in effect controlling the differences in ability in the comparison of UK score. What is left behind in the comparison will be the difference in UK knowledge. This will be the main cause of the variation in the difference/ipsative score among the three HK groups. An inspection of the figures in Table 1 seems to show that UK knowledge is positively related to UK score (see Table 1). In other words, exposure may improve UK knowledge.

In Dolan's original analyses on the four groups (3 HK + 1 UK)(see his Table 3), as the UK group differs from the three HK groups in both UK knowledge and ability, a simple ANOVA on the UK score is problematic and inconclusive. It is uncertain whether the significant F-value is caused by the difference in ability or UK knowledge. The use of the difference/ipsative score does not solve this problem because once the HK score is used in the adjustment (the regression to adjust for sampling differences as used by Dolan), there is an added complication that the UK group differs from the HK groups by HK knowledge as well. Dolan's adjustment for sampling differences will be appropriate only when the UK group has similar HK knowledge as other HK groups. In conclusion, in the examination of the hypothesis of schematic background on recall (the main hypothesis in Dolan's article), it is suggested that Dolan should apply his method (see his Table 3) on the three HK groups rather than on four groups (3 HK + 1 UK) together.

A minor problem was also noted with the reliability coefficients. Dolan reported Cronbach's alpha coefficients for each scale and item. Probably those values for items were erroneously copied from the 'alpha if item deleted ' or 'item-total correlation' information in the computer output (the latter was suggested by an anonymous reviewer). The former reflects the change in alpha including and excluding a particular item, whereas the latter reflects the strength of association between an item and the scale score.

Discussion and Conclusion

The results of the above analysis show that UK exposure can possibly increase specific UK knowledge. The results, however, do not substantiate *Dolan's claim* of 'general improvements in language proficiency' (1994, p.31 abstract and p.35). On the contrary, the reanalyses in this study show that there is no clear trend of substantive improvement in general language proficiency even after nine month study/exposure in UK. If the ultimate and most important goal of exposure programs is improvement of general language proficiency, then the present findings cast serious doubt on the benefits of spending money on exposure programs shorter than nine months.

Admittedly, UK exposure may increase specific UK knowledge, but these benefits have to be evaluated against other, possibly much cheaper or more efficient, alternatives. For example, can the same increase in UK knowledge be gained through

books, lessons, films, or other audio-visual media, in a much shorter time and in a more affordable way? Futhermore, for studies whose main interest is the improvement in general language proficiency, more appropriate instruments which measure a much wider spectrum of language abilities should be used.

It is to be noted that in Dolan's study, the comparisons among HK_{0m}, HK_{3m}, HK_{9m} groups have to be carried out with care because the effect of exposure is confounded with extra English lessons. For example, the HK_{9m} group had advantages in having both greater UK exposure and an extra nine months of intensive English learning from the B. Ed. course they were attending. A possibly stronger research design is to compare the experimental groups with other equivalent groups who attend the same course in HK and are taught by the same instructor. A general language proficiency battery should also be administered throughout various stages of the course. These control groups differ from Dolan's HK_{0m} in that the latter has not attended additional B. Ed. courses. I am not recommending this research design as a subsitute for Dolan's one. Instead, I propose the new design in order to show the confounding effects of UK exposure and extra B. Ed. courses, and hence some possible limitations of Dolan's interpretations of his findings.

In summary, Dolan's study shows a possibility that UK exposure increases subjects' UK specific knowledge and thus improves their recall performance on UK items. But a re-examination of all his analyses reveals no support for *his claim* of the relation between exposure and improvement in general language proficiency. Rather, the above reanalyses show that substantial improvement in language proficiency cannot be gained even after nine months of exposure. From a pragmatic point of view, the benefits of UK exposure are unsubstantiated and dubious.

References

Carroll D. W. (1986). *The psychology of language.* Belmont: Brooks/Cole.

Dolan, T. (1994). Schemata and memory for sentences: The effects of exposure to a British environment on the English of Hong Kong teachers at University in England. *Educational Research Journal, 9,* 31-37.