ELR Journal Vol. 4.

Vol. 4.
1991

ELR JOURNAL

CLASSROOM CONCORDANCING

Vanca Stevens

ELR Journal (New Series) Vol. 4.

CLASSROOM CONCORDANCING

Editors: Tim Johns and Philip King

Centre for English Language Studies
The University of Birmingham
Edgbaston
Birmingham B15 2TT

Contents

CONCORDANCE-BASED VOCABULARY EXERCISES: A VIABLE ALTERNATIVE TO GAP-FILLERS

Vance Stevens

Sultan Qaboos University

INTRODUCTION

GAP FILLER VOCABULARY EXERCISES

Gap-filling is frequently used as a format for vocabulary exercises and tests. Typically, students have been given a list of vocabulary words, and the teacher or materials writer creates a paragraph or two where a subset of these words is contextualized, but the words themselves are blanked out, leaving gaps which the students must fill in with each missing word. In the course of manipulating contexts to accommodate vocabulary (normally, the opposite is done), some of the language surrounding the target vocabulary words may become artificial and contrived.

Despite the popularity of this format, student performance can be unexpectedly poor on this type of exercise. One possible reason is that discoursal clues and constraints on words allowed in a given blank may be well removed from the blank in question, and language learners, particularly at lower levels of proficiency, may not read gap-filling passages as cohesive, continuous text (as suggested in Nunan, 1985). Another reason is that, once a student misplaces one word, then that word is no longer available for placement in its correct position, and a domino effect is set up in which the number of incorrect answers is compounded. This can be discouraging for students; marks on a ten-item exercise, for example, could reach "failure" after misplacing just two words, sending a more negative signal than is deserved.

Exploitation of gap-filler exercises can also be limited. Since license may have been taken with the language in order to fit the context around each vocabulary item, there may be little for a teacher to discuss in following up such exercises, apart from the meaning of the words that should have been put in the blanks. On occasion, it may be worth pointing out to students discourse elements in the passage determining certain word choices, or the subject matter of the reading may be of passing interest, but a text created solely for the purpose of providing context for a certain set of vocabulary items may be neither relevant to the students nor rich in natural language.

CONCORDANCE-BASED VOCABULARY EXERCISES

A computer-based concordance program will find each instance of a search string (e.g. a vocabulary item) in a text database. Key-word in context (KWIC) is the most common way of displaying such concordances; that is, each time the program encounters an instance of the string it is looking for in the database, it prints that string in the middle of the screen, along with enough of the context either side (as it appears in the text) to fill up the screen. A sample raw concordance output is shown in Appendix 1.

Concordance-based vocabulary exercises can be derived from any text database deemed suitable; however, Stevens (forthcoming) discusses the advantages in utilizing student course books as the database. Concordances on the words appearing in the vocabulary lists in course books ought to turn up several examples for each target word (if not, then the vocabulary list probably should be revised). Stevens (1988 and forthcoming) maintains that vocabulary exercises derived in this way contain items that are authentic, since they present contexts from language created for a communicative purpose, and relevant, since the contexts are ones that the students are at that moment studying. (See Appendix 2 for a sample concordance-based exercise; note how the items were derived from the concordance output shown in Appendix 1.)

The problem of one misplaced word displacing another is attenuated through use of concordance-based exercises. Concordance-based exercises are similar to gap-filler ones in that students replenish gaps from a menu of words arrayed at the top of the page; however, concordance-based exercises present multiple contexts - since the items in a concordance-based exercise are excerpted from raw concordance output (block-copied right into the exercise using a word processor), the gaps occur where the target words, centered in a column one above the other, have been deleted from each item. Because they have several contexts with which to double-check the suitability of each word placed, students may do better on such exercises than with their gap-filler counterparts, while still accomplishing the purpose of the exercise; i.e. reinforcement of vocabulary.

A possible counter-argument is that students shouldn't do as well on such exercises because the items drawn directly from raw concordance output are truncated at either end, and students are therefore presented with fragmented text. I contend that this is actually an advantage lending itself to fruitful exploitation in subsequent class periods, because students can extrapolate the information missing at either end of the concordance output, and the more they extrapolate, the more they can recall and discuss the original context. The ability (or willingness) to extrapolate holistically from fragmentary evidence is a vital skill sadly lacking, even at the highest levels, in students passing through many educational systems; still, some teachers question on purely intuitive grounds the desirability of students coping with truncated text, thus further postponing the students' ever coming to grips with real-world data.

Johns (1988: 9) has noted that concordancing "tends to divide language teachers into two

camps", those who are enthusiastic about the concept, and those who are "puzzled by it <and> have failed to see that it could be of any use to a learner ... This division has little to do with language teachers' alleged fear of computer technology, and a great deal to do with underlying assumptions about the nature of language learning and the role of the teacher in that process." Johns suggests that an appropriate role for a language teacher is 'director of research', an idea which he expands (Johns, 1991) into the concept of data-driven learning, in which learners' competence is developed through exposure to samples of linguistic performance. In Johns' words, "we simply provide the evidence needed to answer the learner's questions, and rely on the learner's intelligence to find answers." (p. 2)

Higgins (1983) characterizes the occupants of Johns' two camps as pedagogues and magisters, respectively. Pedagogues stand ready with facilitative information supplied to students on demand, when the students feel ready. Magisters, on the other hand, take charge of the learning process themselves and control delivery of information to students. This leads to an unexpected consequence, in that an intelligent tutor can breed the opposite quality in students and vice versa:

> the machine's lack of intelligence shows us things we might never have noticed for ourselves and awakens intelligence and imagination people who have had little chance to develop them before. This is in contrast to those approaches to language teaching, regrettably common, which assume a teacher who is both proficient in the subject matter and intelligent about deciding how to present it, while also assuming a learner who has no proficiency and no intelligence. (Higgins, 1988:51)

At the root of this division between the two camps is the question: to whom are decisions regarding what-is-to-be-learned-and-when most appropriately left? To learners or their teachers? Surely, input from both camps is relevant, and the most salient flaws in the arguments emanating from either occur where staunch advocates of one group neglect the needs or guidance of the other.

THE RESEARCH QUESTION

Detractors of concordance-based exercises may doubt that students, left to their own devices, are able to make the mental leaps required to learn for themselves a system as complex as that of a living language. The present study attempts to counter this notion by disproving the null hypothesis that exercises drilling the same vocabulary in gap-filler and concordance-based formats can be solved equally well by language learners, and furthermore by hypothesizing that differences in performance will favor the concordance-based exercises.

SUBJECTS

The subjects were first year male and female Arab university science students enrolled in an

English-for-physics course given by the Language Centre at Sultan Qaboos University in Oman. This course was meant to support a first-year physics course taken simultaneously from the College of Science. The subjects were grouped in two Language Centre classes, and their teachers carried out the experiment during regularly scheduled class times.

PROTOCOL

The experiment was run twice in each of the two participating classes; that is, the study was conducted in Sessions 1 and 2 for Groups 1 and 2 (the treatment for one group being simply replicated in the other). Groups 1 and 2 were in turn divided into halves, each of whom worked one concordance-based and one gap-filler vocabulary practice exercise per experimental Session. In an attempt to mitigate the effects of unfamiliarity with one of the formats, all students were given practice with concordance-based exercises prior to the experiment.

Session 1 treatment

During a single scheduled class period, gap-filler and concordance-based exercises testing identical vocabulary items were handed out to alternately seated students (alternate distribution prevented students copying each other's work and interjected an element of randomness in the assignment of treatment). Students were given a set time to work the exercises. After these were collected, a second set of papers, each testing identical vocabulary items (but different from the first set), was distributed in such a way that each student who received a gap-filler exercise in the first round got a concordance-based one in the second, and visa versa. The students were given the same amount of time as before to complete these exercises.

Session 2 treatment

The experiment was repeated at a later date during a second scheduled class period in exactly the same way, but with still different vocabulary items.

THE INSTRUMENTS

Eight instruments were devised for the experiment: four concordance-based and four gap-filler; that is, two of each for any one experimental session (one gap-filler and one concordance-based exercise each testing one set of vocabulary, and one each testing a second set of vocabulary). In addition, practice concordance-based exercises were needed to familiarize students with the concordance-based format prior to their participating in the experiment.

Each instrument had ten items; that is, each gap-filler exercise had ten words blanked out, and each concordance-based exercise had ten sets of contexts for the same vocabulary items as on the corresponding gap-filler. Each exercise had almost identical directions, with the same 12

words (10 correct answers and 2 distractors) arrayed in identical order at the top of each gap-filler and its concordance-based counterpart.

At the time of this experiment, the subjects were using an English-for-physics reading textbook with lists of vocabulary items at the end of each unit. For some units, the textbook contained gap-filler exercises for vocabulary practice, and the two reading teachers involved in the experiment occasionally prepared their own gap-filler exercises to drill that vocabulary as part of their routine in class. Furthermore, the students had doubtlessly encountered the gap-filler format numerous times throughout their student careers.

Since the purpose of this experiment was to test an alternative to the kind of gap-filler exercise created by teachers in the normal course of teaching a class, it would have been ideal if the teachers concerned could have first created their gap-filler exercises, after which concordance-based exercises could have been created testing the vocabulary appearing on the spontaneously created gap-filler exercises. However, this turned out to be impractical because concordances of words on the lists at the back of each unit did not all produce a sufficient number of robust items. Therefore, in order that all the instruments used would be of fairly consistent validity, vocabulary items eligible for the gap-filler exercises had to be restricted to those which also produced suitable concordance material.

A concordance-based item was considered to be robust only when it could be derived from raw output giving at least three, and preferably four, appropriately contextualized occurrences of that item in the reading text. The number of contexts per question could of course become a variable in future research, but for this experiment, that number was set at three or four contexts, (although for lack of truly robust items, one item with only two contexts was included on each concordance-based instrument for Session 2).

Even where raw concordance output produced sufficient occurrences of a word, not all lines-of-context were appropriately contextualized. For instance, definitions (where the context would define the word) were ruled out. Also, care had to be taken that the answer to one question did not appear in the context for another elsewhere on the same instrument. Furthermore, the target word in a context used in the instrument had to appear with the same meaning and part of speech as it was given in the word list. Finally, the context had to be meaningful; that is, it had to contain somewhere in it some clue, however small, to assist students in placing the target word in that context.

Wherever these ground rules were violated in the raw concordance output, that line-of-context was eliminated from consideration for the final instruments, and vocabulary items for which there remained only one or two robust lines-of-context were also eliminated from consideration (with the exception of one item each on the Session 2 instruments, as noted above). In the many instances where items had ample robust contexts, three or four of the most illustrative were selected to appear on the instruments.

Once it was discovered what words on the vocabulary list for a given unit produced robust items for the concordance-based instruments, the subject teachers were asked to create gap-filler exercises using those words. However, due to time constraints, they were only able to do this themselves for the instruments used in Session 1. The researcher therefore had to produce the gap-filler materials used in Session 2. However, the researcher, having previously taught a similar course, produced exercises which in the opinion of the subject teachers were quite similar to ones they would have produced, so the integrity of the experiment was not compromised by this expediency.

Subject matter for three of the gap-filler exercises was on the topic of physics, and on aspects of that subject deemed accessible to the students (in fact, one topic in each experimental session was directly related to topics the students had studied in the unit concerned); however, one of the English teachers created a gap-filler based on the proverb of the goose who laid the golden egg. This was considered acceptable, since it was in fact a "gap-filler exercise created by a teacher in the normal course of teaching a class."

The experimenter and two teachers were of the opinion that the golden egg story was the easiest of the four gap-filler instruments, and that the other gap-fillers were pitched at different reading levels. No attempt was made to control reading level on these instruments, except that all three of us wrote at a level that we felt the students should be able to comprehend, according to our experience with them. It should be remembered however that the first priority in the creation of gap-filler exercises is to think of some context that will accommodate the vocabulary at hand; if this can be done with a subject familiar to the students and at a level all can understand, then this is a reflection of the skill of the materials writer - but "in the normal course of teaching a class," such considerations can easily become secondary to the challenge of contextualizing the vocabulary.

The contexts for the concordance-based exercises should have been familiar to the students since they were taken straight from their textbooks (presuming, of course, that the students had read their books; their teachers were of the opinion that in general, they had). Session 1 occurred at the time that the subjects had finished unit 5 in their course book, and so the instruments were created from units 4 and 5 vocabulary lists, with the concordance database being units 1 through 5 in the book. Session 2 occurred just as the students were completing unit 6; therefore, the instruments were created from the unit 6 vocabulary list (divided into two parts), with the concordance database this time being units 1 through 6 in the book.

The first step in preparation of the concordance-based exercises was to create an ASCII version of the student course book. This was accomplished simply by running a program that stripped off the high bit from each byte in the word-processed version. Fortunately, the book had been created on a word processor, and the soft copy version was available to the researcher.

Once that initial step had been accomplished, each concordance-based exercise took about an hour to create from start to finish. Steps included (1) writing and running a batch file to run the

concordances on the vocabulary from each unit (and making coffee while the computer went about that task); (2) concatenating the concordance output files into one large word processor file; (3) zapping (i.e. eliminating) all non-robust lines-of-context, items with fewer than two contexts, and extraneous material with the word-processor's line-delete function; (4) scrambling the usable items with block moves; (5) loading in the rubric from a previous file and making minor changes; (6) blanking out the target words and replacing these with boxes, also by block moves; and (7) voila! another concordance-based exercise. It took at least as much time (or longer) and twice the brain-strain to make the gap-filler exercises, allowing no time for making coffee.

Block-removing the concordanced words from four or five lines-of-context so that the same word could be replaced in all four contexts meant that any inflections occurring in the concordanced discourse had to be left in. The author has worked with concordance-based exercises where the inflections are removed, so that the students must provide the correct form of the target word, but for this to have been viable in the present experiment, it would have been necessary to use subjects experienced with the technique. Therefore, inflections were left in the concordance-based instruments where they stood, and to make the gap-fillers commensurate with their counterparts, inflections were retained there also. No attempt whatsoever was made to select lines-of-context in the concordance-based instruments or to manipulate syntax in the gap-fillers to either increase or minimize the use of inflected forms.

RESULTS

The results are given in Tables 1 through 6. Tables 1 and 2 show results from Session 1, in which each subject in Groups 1 and 2 either worked GF4 and CB5, or GF5 and CB4; that is, the gap-filler and concordance-based treatments utilizing vocabulary taken from units 4 and 5 of their English-for-physics textbook. Raw scores are given along with a statistical analysis of each group's overall performance on the gap-filler exercise relative to its performance on the concordance-based one. Tables 3 and 4 give similar information for Session 2, in which the same two groups of students repeated the experiment, this time with exercises derived from vocabulary sets A and B in unit 6; that is, each subject worked either GF6a and CB6b, or GF6b and CB6a.

Each table includes two analyses of the data. Because the data occur in non-continuous integers, they violate the assumption of normality which is a condition of using the t-test. Therefore, a non-parametric measure, the Wilcoxon Matched-pairs Signed-ranks Test, was used to test the null hypothesis that subject performance on both gap-filler and concordance-based exercises was essentially the same. However, as the t-test is considered robust even with certain deviations from its assumptions, it served as a check on the results from the non-parametric measures, which in all cases were confirmed.

Results from Session 1 were inconclusive. In general, students in Group 1 did better on the concordance-based exercises, but not significantly so. Results for subjects in Group 2 were erratic, with subjects doing significantly better on CB4 than on GF5, yet significantly better on GF4 than on CB5; thus one set of results favored the concordance-based exercise and the other the gap-filler. Session 2 results clearly favored the concordance-based exercises, with both groups of students performing significantly better on the these than on the gap-filler ones for all four pairings (with p values ranging from .0464 to .0051 on the non-parametric measures).

To test whether outliers in any one group may have skewed that group's scores, the two groups were treated as one and the analysis repeated for each session (Tables 5 and 6). The results show that, for Session 1, all subjects in both groups combined did significantly better on CB4 than on GF5 whereas there was no significant difference in performance on GF4 over CB5 when both groups were combined. Thus the significant finding in Session 1 favoring the gap-filler exercise above was not supported, whereas that for the concordance-based exercise was. Findings for Session 2 favoring the concordance-based exercises were confirmed in this check on the analysis.

DISCUSSION OF RESULTS

There are many factors to consider in comparing two exercise types. In many respects, the comparison is similar to that between apples and oranges; that is, without pinning down and controlling for a welter of variables (e.g. readability, cloze frequency, number of words per instrument, syntactic and lexical density, familiarity with exercise type, and varying nature of the subject matter on each instrument) it is impossible to say that one type of exercise will always be easier or more difficult than another.

One factor that was anticipated and dealt with was familiarity with exercise format. Whereas the gap-filling exercise format is widely used throughout our students' careers, our student subjects had no idea what a concordance was, and so had to be familiarized with the concordance-based vocabulary exercise concept. Both groups of students therefore were given a pilot concordance-based exercise to work prior to the experiment. Group 1 worked with concordance-based exercises for two class sessions prior to doing the experiment; Group 2 worked with concordance-based exercises only the hour prior to doing the experimental treatment. This could partially explain the ambivalent results from Session 1 for Group 2.

Several factors may have contributed to the significantly better performance of students on the concordance-based exercises in Session 2. One factor would of course be increased familiarity, and the results are consistent with the possibility that the concordance-based exercises were easier for the students once they had gained some experience with them.

Another factor favoring the concordance-based exercises may have been relative difficulty of the gap-filler exercises. Although no assessment was made of degree of difficulty of the gap-

fillers, it could have been that the ones used for Session 2 were more difficult than those used in Session 1. (On the other hand, the results favoring GF4 in Session 1 may have been due in large part to the fact that this particular gap-filler, about the goose that laid the golden egg, was relatively easy for the students to read.)

Still another factor might have been familiarity with the parts of the textbook from which the concordances were derived. If this was indeed a factor, the students may have been able to provide some answers because they recalled from memory the missing word in one or more of the three or four contexts. In this event, it could have been familiarity with the text itself, as opposed to familiarity with the exercise format, that allowed the students to answer significantly more questions correctly on one or more concordance-based instruments.

Although the above factors may have all played some part in the results, it is hoped (and considered possible) that the generally better performance (significantly better in Session 2) of each experimental group on the concordance-based exercises was due, at least in part, to the fact that the subjects were given more contexts with which to narrow down their choices of what word to fit in what blank. It is certainly clear that the concordance-based format did not interfere much with the subjects' ability to carry out this task, except perhaps during a brief familiarization phase.

CONCLUSION

Although it would be inappropriate to say from these data that concordance-based exercises derived from a text base familiar to the students are superior to the kind of gap-fillers commonly created by teachers as vocabulary exercises, it may be concluded that they can indeed be more easily solved. This is not to say that concordance-based exercises of the type investigated here are any less challenging than gap-fillers; concordance-based exercises require students to extrapolate beyond fragments of discourse and to puzzle out what word several such fragments have in common. When the corpus used is the students' own textbooks, "easiness" deriving from recall from memory of some of the lines-of-context need not be considered a problem - if the purpose of the exercise is to reinforce the vocabulary, as opposed to testing, and if the proclivity of the teacher is to engender a sense of confidence and well-being in the students with regard to the language under study, then concordance-based exercises are viable alternatives to gap-filler ones.

What has been established in this study is that students' being confronted with exercises reflecting the truncated demi-contexts typical of concordance output does not seem to be a hindrance to their discerning the word missing from those contexts. If anything, having multiple if disjunct contexts helps them more in settling on a correct word than do the clues inherent in a passage of discourse with the same words missing.

Despite the many other factors involved, it is assumed that increased familiarity with the format played some part in the improvement, for both groups, in scores on the concordance-based exercises compared to those on the gap-filler when the experiment was run the second time. Although there was no time to continue the experiment beyond Session 2, it would be interesting to see if the results would maintain through repeated sessions, or if the much higher concordance-based scores for the second session were aberrations influenced by the many other variables alluded to above.

ACKNOWLEDGMENTS

The quality of this experiment was greatly enhanced by the gracious assistance of Tom Cobb and Marlise Horst, reading teachers, who carried out the experimental treatments in their classes. Thanks also to Tahsin Erkan Ture, Dept. of Maths and Computing at Sultan Qaboos University, Oman, whose advice on experimental design and help with subsequent analysis was invaluable.

REFERENCES

Higgins, J. (1983) Can computers teach?, *CALICO Journal*, 1:2, 4-6.

Higgins, J. (1988) *Language, learners and computers*. London and New York: Longman.

Johns, T. (1988) Whence and whither classroom concordancing? In Bongaerts et al. (Eds.) *Computer applications in language learning*. Foris.

Johns, T. (1991) Should you be persuaded? - Two samples of data-driven learning materials. This volume

Nunan, D. (1985) Content familiarity and the perception of textual relationships in second language reading. *RELC Journal* 16, 43-50.

Stevens, V. (Forthcoming) Classroom concordancing: Vocabulary materials derived from relevant, authentic text. *English for Special Purposes Journal.*

Stevens, V. (1988) Studying vocabulary using concordances on microcomputers. *TESOL Newsletter* 22, 3:27.

Table 1

Session 1, Group 1:

Gap-Filler and Concordance-Based Scores (out of 10), Units 4 & 5

St. #	GF4	GF5	CB4	CB5
G1-1		6	7	
G1-2		7	10	
G1-3	8			5
G1-4	10			10
G1-5				
G1-6	8			9
G1-7	8			5
G1-8		9	8	
G1-9		8	10	
G1-10	7			6
G1-11		8	10	
G1-12	10			10
G1-13	5			7
G1-14				
G1-15	7			9
G1-16		7	10	
G1-17		8	9	
G1-18	6			9
G1-19		9	7	
G1-20		10	10	
G1-21	5			6
G1-22				
G1-23		6	3	
G1-24	8			10
G1-25		10	9	
n size	11.00	11.00	11.00	11.00
avg	7.45	8.00	8.45	7.82
st.dev.	1.62	1.35	2.06	1.95
var	2.61	1.82	4.25	3.79

Comparing Group 1, GF5 with CB4

Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -.7645; p = .4446; no significant difference

Paired samples t-test:
2-tailed, 10 d.f.; t = -.75; p = .472; no significant difference

Comparing Group 1, GF4 with CB5

Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -.5331; p = .5940; no significant difference

Paired samples t-test:
2-tailed, 10 d.f.; t = -.60; p = .563; no significant difference

Table 2

Session 1, Group 2:

Gap-Filler and Concordance-Based Scores (out of 10), Units 4 & 5

St. #	GF4	GF5	CB4	CB5
G2-1				
G2-2	8			2
G2-3		2	7	
G2-4	9			6
G2-5		10	8	
G2-6	9			10
G2-7		5	10	
G2-8		3	8	
G2-9		3	7	
G2-10		5	10	
G2-11		7	10	
G2-12		7	10	
G2-13		3	9	
G2-14		7	8	
G2-15	10			5
G2-16	8			3
G2-17	10			9
G2-18	6			4
G2-19		7	10	
G2-20	9			8
G2-21				
G2-22				
G2-23	9			6
G2-24		1	3	
G2-25	7			7
G2-26		4	6	
G2-27	5			4
G2-28	3			5
G2-29	9			9
n size	13.00	13.00	13.00	13.00
avg	7.85	4.92	8.15	6.00
st.dev.	1.99	2.46	1.99	2.39
var	3.98	6.07	3.98	5.69

Comparing Group 2, GF5 with CB4
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -2.9701; p = .0030; * significantly better on CB4

Paired samples t-test:
2-tailed, 12 d.f.; t = -5.38; p = .000; * signficance confirmed

Comparing Group 2, GF4 with CB5
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -2.2228; p = .0262; * significantly better on GF4

Paired samples t-test:

2-tailed, 12 d.f.; t = 2.72; p = .018; * signficance confirmed

Table 3

Session 2, Group 1:

Gap-Filler and Concordance-Based Scores (out of 10), Unit 6

St. #	GF6a	GF6b	CB6a	CB6b
G1-1	4			10
G1-2	4			10
G1-3	6			7
G1-4				
G1-5	3			10
G1-6		9	8	
G1-7		8	7	
G1-8		7	10	
G1-9	5			10
G1-10		7	9	
G1-11	7			10
G1-12		6	10	
G1-13		7	9	
G1-14				
G1-15	5			4
G1-16	3			10
G1-17	4			10
G1-18	5			8
G1-19	8			10
G1-20		6	10	
G1-21		4	10	
G1-22				
G1-23		10	10	
G1-24				
G1-25		7	10	
n size	11.00	10.00	10.00	11.00
avg	4.91	7.10	9.30	9.00
st.dev.	1.50	1.58	1.00	1.86
var	2.26	2.49	1.01	3.45

Comparing Group 1, GF6a with CB6b
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -2.8007; p = .0051; * significantly better on CB6b

Paired samples t-test:
2-tailed, 10 d.f.; t = -5.10; p = .000; * signficance confirmed

Comparing Group 1, GF6b with CB6a
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -2.3102; p = .0209; * significantly better on CB6a

Paired samples t-test:
2-tailed, 9 d.f.; t = -3.03; p = .014; * signficance confirmed

Table 4

Session 2, Group 2

Gap-Filler and Concordance-Based Scores (out of 10), Unit 6

St. #	GF6a	GF6b	CB6a	CB6b
G2-1				
G2-2		4	10	
G2-3	3			6
G2-4	3			10
G2-5	4			10
G2-6	10			10
G2-7				
G2-8				
G2-9				
G2-10	6			9
G2-11	6			9
G2-12		7	10	
G2-13	0			4
G2-14		6	8	
G2-15		4	6	
G2-16		7	9	
G2-17		7	10	
G2-18		3	6	
G2-19				
G2-20				
G2-21				
G2-22		3	3	
G2-23		7	8	
G2-24				
G2-25				
G2-26	2			3
G2-27				
G2-28				
G2-29	8			10
n size	9.00	9.00	9.00	9.00
avg	4.67	5.33	7.78	7.89
st.dev.	2.94	1.70	2.25	2.64
var	8.67	2.89	5.06	6.99

Comparing Group 2, GF6a with CB6b
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -2.5205; p = .0117; * significantly better on CB6b

Paired samples t-test:
2-tailed, 8 d.f.; t = -4.35; p = .002; * signficance confirmed

Comparing Group 2, GF6b with CB6a
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -1.9917; p = .0464; * significantly better on CB6a

Paired samples t-test:
2-tailed, 6 d.f.; t = -3.00; p = .024; * signficance confirmed

Table 5

Session 1, Groups 1 and 2

Combined Statistics

Comparing Groups 1 and 2, GF5 with CB4
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -3.0567; p = .0022; * significantly better on CB4

Paired samples t-test:
2-tailed, 23 d.f.; t = -3.85; p = .001; * signficance confirmed

Comparing Groups 1 and 2, GF4 with CB5
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -1.4186; p = .1560; no significan difference

Paired samples t-test:
2-tailed, 23 d.f.; t = 1.65; p = .113; no significant difference

Table 6

Session 2, Groups 1 and 2

Combined Statistics

Comparing Groups 1 and 2, GF6a with CB6b
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -3.7425; p = .0002; * significantly better on CB6b

Paired samples t-test:
2-tailed, 19 d.f.; t = -6.75; p = .000; * signficance confirmed

Comparing Groups 1 and 2, GF6b with CB6a
Wilcox Matched-pairs Signed-ranks Test:
2-tailed; Z = -3.0670; p = .0022; * significantly better on CB6a

Paired samples t-test:
2-tailed, 16 d.f.; t = -4.34; p = .001; * signficance confirmed