

# **User Performance With Audio: The Effect Of Users' Cognitive Styles**

**David John**

School of Design, Engineering & Computing

Bournemouth University

djohn@bournemouth.ac.uk

**Professor Anthony Boucouvalas**

School of Design, Engineering & Computing

Bournemouth University

tboucouv@bournemouth.ac.uk

## **ABSTRACT**

This work is concerned with the issue of user performance with audio interfaces. We examine how people perform when assigned with simple purely audio tasks. We define 'user performance' as the ratio of correct score divided by the time to complete the task. A test was devised to examine how the user cognitive style classification is related to user performance when dealing with audio tasks. Performance was significantly different between all tasks except for two cases. Subjects performed best where the level of cognitive demands placed on them were low and prefer questions that use 'Speech' rather than 'Non-Speech'. Cognitive style was not found to be a significant factor influencing performance where all subjects performed well, but the Verbal-Imagery dimension was found to be a significant factor for the more complex tasks.

This paper reports on an experiment that examines how subjects perform in a number of simple tasks using audio. Performance was measured in terms of time taken to answer a series of questions and the number of correct answers scored. There were two main aims for conducting the experiment. The first aim was to identify the relative performance of subjects in all types of task. The second aim was to determine whether performance differences between the subjects was attributable to their cognitive style. The Cognitive Styles Analysis (CSA) classification of individuals (Riding 1991) was used to classify the cognitive style subjects. The results of the experiment have implications for the interface design of adaptive and non-adaptive computer systems.

The experiment was part of the Telecare Companion project (John, Jerrams-Smith et al. 1998) that was investigating how multimedia computer systems could automatically adapt aspects of the interface to suit the needs of individual users, and the nature of the task. The emphasis of the Telecare Companion project is on identifying the influence of the emerging multimedia technologies, including how different media can be used in the interface, and what different adaptive strategies can be followed by using different media.

Computer systems can automatically adapt to individual users by determining the most appropriate manner to interact with users based on what it 'knows' about the individual users using information held in the *user models* (Benyon and Murray 1993; Brusilovsky 1996; Elsom-Cook 1993; Self 1988). Using techniques such as the *explanation variants* technique (Beaumont 1994; Bocker, Hohl et al. 1990; Hohl, Bocker et al. 1996) a computer system selects the piece of information from a number of different versions of the same information that is most appropriate for the individual.

By analysing the results of the audio experiment it may be possible to derive general recommendations for developers of multimedia computer systems by identifying audio tasks for high performance. Also the results would determine whether there is any value in preparing alternative versions of the same information that use audio to suit different cognitive styles. Where cognitive style is found to effect the performance of subjects there is a potential benefit in adapting computer systems to suit the cognitive style of users.

This paper first describes how the audio experiment was designed. Next the results are presented, including the difference in subjects' performance between the tasks and the difference in performance between the cognitive style groups within each task. Finally some of the issues identified by analysing the results are discussed in more detail.

## DESIGN OF THE AUDIO EXPERIMENT

The experiment was testing subjects' perception of audio information that was being presented by a multimedia computer system. Subjects were first presented with audio information and their understanding of the information was tested using a multimedia interface. Questions were asked using audio but subjects entered answers by clicking on possible answers represented by either text or images.

In order to examine differences in the performance of subjects using audio a wide range of tasks were used. The main division in the types of audio questions asked in the experiment was between the 'Speech' ('S') questions that use language, and the 'Non-Speech' ('NS') questions that do not use language.

The 'S' questions involved simple comparison questions, comprehension questions and mathematical questions. The comparison questions required subjects to consider two words and determine whether they were the same type or colour. There were two categories of words used in the experiment; concepts and objects. The *comparison of concepts questions* contained words that were not objects, such as 'Rugby' and 'Tennis', while the *comparison of objects questions* described the appearance of objects such as 'mug' and 'cup'. In the *comprehension questions* subjects first heard a set of procedural instructions, such as directions and distances between landmarks in a town, and were then asked a set of multiple-choice questions about details of the information, such as how the landmarks relate to each other geographically. The *mathematical questions* asked subjects to perform simple addition, multiplication and division calculations.

The 'NS' questions examined how individuals perceive sound that does not contain words. Tasks were prepared to examine the major phenomenological aspects that effect the perception of sound. Aspects of the perception of sound that were considered included *pitch* (the frequency of the sound wave), *loudness* (or intensity), *timbre* (the difference between two different instruments playing at the same pitch and frequency), *consonance* and *dissonance* (two or more tones played simultaneously), and *location* (where the sound is coming from) (Gross 1992; Goldstein 1989; Handel 1993). Subjects' perception of timbre, consonance and dissonance were examined by testing their ability to recognise musical instruments and sound effects. A number of musical instruments were played to subjects with consistent pitch, loudness and location and then subjects were then asked to identify individual instruments, or two instruments playing together by selecting the answer form a multiple-choice list.

Subjects were also asked to identify sound effects by selecting the image that represented the source of the sound.

## RESULTS OF THE EXPERIMENT

The experiment was performed using fifty subjects who were selected at random from the staff and students of Bournemouth University using a multimedia PC. The performance of each subject was measured by recording the number of questions they answered correctly and the amount of time in seconds that they took to answer each question. The measurement of duration recorded subjects' reaction time in answering questions and not the time spent listening to the questions.

In order to compare the relative performance of subjects between each task a number of calculations had to be performed on the results. Scores were calculated as the percentage of questions answered correctly and duration figures were calculated as *the average time* taken to answer one question. Viewing score or duration alone may give a distorted view of overall performance as a subject who achieved a high score may have taken a long time to answer the questions, and a subject who completed the task in a short time may have a low score. Therefore, a performance measure has been devised that combines the score and duration to provide an overall measure of performance. For each subject scores are divided by duration to form a performance ratio. For subjects to have a high performance ratio they have to score high in a short duration. Where all subjects scored a similar amount differences in the performance of subjects are magnified by dividing the scores by the duration. However, where there is a wider variation between the scores of subjects the effect is lessened by dividing the scores by the duration.

The audio questions were divided into 'S' and 'NS' type questions and so it was important to compare the relative performance of subjects in both types of questions. In order to assess how subjects performed giving equal weight to both the 'S' and 'NS' questions an average was calculated. For all subjects the following calculations were performed:

1. 'S' Questions Score =  $\sum S1/N1$  %
2. 'S' Questions Duration =  $\sum D1/N1$
3. 'S' Questions Performance Ratio =  $\sum S1/D1$
4. 'NS' Questions Score =  $\sum S2/N2$  %
5. 'NS' Questions Duration =  $\sum D2/N2$
6. 'NS' Questions Performance Ratio =  $\sum S2/D2$
7. Average Score =  $\sum ((S1/N1) + (S2/N2))/2$  (%)
8. Average Duration =  $\sum ((D1/N1) + (D2/N2))/2$
9. Average Performance Ratio = Average Score / Average Duration

Where;

S1 is the sum of correctly answered 'S' questions  
S2 is the sum of correctly answered 'NS' questions  
D1 is the time taken to complete all the 'S' questions  
D2 is the time taken to complete all the 'NS' questions  
N1 is the number of 'S' questions  
N2 is the number of 'NS' questions

## Relative performance between the tasks

Comparing the relative performance of subjects between the tasks would identify the types of task in which subjects perform relatively well using audio and the types of task in which subjects perform relatively badly using audio. While it is not possible to determine whether audio is more or less suited than other media for each type of task these results will provide developers of multimedia computer systems an indication of the expected performance of users with each type of task when using audio.

The mean performance ratios for all subjects in all of the audio tasks are shown in the following table:

**Table 1**  
Mean Performance Ratios For All Audio Experiment Tasks

Table 1 shows the mean performance ratios, number of subjects and standard deviations for all subjects for all types of question. The "All Questions" row shows the results for all types of question. The next two rows show the results split into the 'S' questions and 'NS' questions, while the next row shows the average between them. The next seven rows show the performance results for the individual tasks.

Comparing all subjects mean performance ratios for the 'S' questions and 'NS' questions using a t-test for related samples (Howell 1995) show that the two sets of results are significantly different from each other ( $t = -10.99$ ,  $df = 49$ ,  $p = 0.00$ ). Therefore subjects performed significantly better with the 'S' questions than the 'NS' questions, achieving a mean performance ratio of 11.12 for 'S' questions compared to 7.54 for 'NS' questions.

The relative performance of subjects in each type of task can be ranked as follows (in descending order):

1. Comparison of objects questions
2. Comparison of concepts questions
3. Identification of sound effects questions

4. Mathematics questions
5. Identification of one musical instrument questions
6. Procedural questions
7. Identification of two musical instrument questions

The ranking order has to be amended after comparing each set of performance ratios against each other using *t*-tests. It is not possible to rank two types of task if the results of each task are not significantly different from each other. The results of the top three types of question were found to be significantly different from each other, i.e. the results of the *comparison of objects questions* were significantly different from the results of the *comparison of concepts questions* ( $t = 6.12, df = 49, p = 0.00$ ), and the results of the *comparison of concepts* were significantly different from *the identification of sound effects questions* ( $t = 2.42, df = 49, p = 0.02$ ). The results of *the identification of sound effects questions* were not significantly different from the results of the *mathematics questions* ( $t = -1.48, df = 49, p = 0.15$ ). The results of the *mathematics questions* were significantly different from the results of the *identification of single instrument questions* ( $t = 2.03, df = 49, p = 0.05$ ). The results of *the identification of single instrument questions* were not significantly different from the results of the *procedural questions* ( $t = -1.45, df = 49, p = 0.15$ ). The results of the *procedural questions* were significantly different from the results the *identification of two musical instrument questions* ( $t = -10.62, df = 49, p = 0.00$ ). Therefore there is no distinction between the results of the *identification of sound effects questions* and the results of the *mathematics questions* tasks ranked 3 and 4, and there is no distinction between the results of the *identification of single instruments questions* and the *procedural questions* that are ranked 5 and 6.

The relative performance of subjects in each type of task can be revised as follows (in descending order):

1. Comparison of objects questions
2. Comparison of concepts questions
3. Identification of sound effects questions **and** Mathematics questions
4. Identification of one musical instrument questions **and** Procedural questions
5. Identification of two musical instrument questions

The implications of the relative ranking of tasks are discussed in Section 4.

## **Results by cognitive style**

Section 3.1 established that the performance of subjects was significantly different between all of the tasks except in two cases. This section attempts to determine whether the performance of subjects within each task varies as a function of the subjects' cognitive style. Where the performance of subjects in different cognitive style groups are significantly different then cognitive style is a major determinant of performance.

### **Cognitive styles analysis**

A cognitive style is the consistent underlying method and thinking and perceiving that affects the way in which an individual perceives and responds to events and ideas (Tennant 1988; Riding and Cheema 1991). Cognitive styles appear to be constant throughout an individual's lifetime and there may be a physical basis for them (Riding 1994; Riding, R. J., Glass et al. 1997).

The CSA package (Riding 1991) is a computerised test that calculates a person's cognitive style as a function of two independent dimensions; the Wholist-Analytic dimension and the Verbal-Imagery dimension. The Wholist-Analytic dimension determines whether people understand situations as a whole, or as see things in parts. Wholists tend to understand situations as a whole and may not distinguish between the parts while Analyticians tend to see the situation as a collection of parts instead of the overall context. The Verbal-Imagery dimension determines whether individuals are out-going and verbal, or more introverted and think in terms of mental pictures or images (Riding 1994). The Wholist-Analytic dimension affects the way individuals make sense of the structure of information (the sequence in which pieces of information are presented) and the Verbal-Imagery dimension affects the way individuals understand the mode of presentation (the mix of text and images) (Riding 1996).

Three categories are defined along the Wholist-Analytic dimension (Wholist, Intermediate and Analytic) and three categories are defined along the Verbal-Imagery dimension (Verbaliser, Bimodal and Verbaliser). The two dimensions are independent of each other but combined form a grid with nine sections (Riding 1991).

An individual's cognitive style is automatically calculated by the CSA program from their responses in a number of simple questions (number of correct answers and response time). There are a number of questions designed to represent all cognitive styles. From the relative performance in each type of question the individual's position along each dimension can be calculated. The individual's position along the Wholist-Analytic dimension is assessed using geometrical shapes. Wholists perform relatively better at identifying whether pairs of complex geometric figures are the same or different, while Analytics perform relatively better at identifying whether simple geometrical shapes are contained in more complex geometrical figures. The questions that assess the individual's position along the Verbal-Imagery dimension use a number of statements that have to be judged true or false. Verbalisers perform relatively better at statements that contain information about conceptual categories while Imagers perform relatively better at statements that describe the appearance of objects (Riding and Staley 1998).

## **Cognitive Style Of Subjects**

The fifty subjects who undertook the audio experiment also undertook the CSA test. In order to simplify the analysis of the experiment, as performed in other experiments (Riding and Sadler-Smith 1992), the cognitive style dimensions were divided into two halves not three (Figure 1). This gave a fairly even division of subjects between the two halves of the Wholist-Analytic dimension (23 Wholist subjects and 27 Analytic subjects), with a slightly more uneven division of subjects between the two halves of the Verbal-Imagery dimension (21 Verbaliser subjects and 29 Imager subjects).

**Figure 1.** Cognitive Style Of Subjects Scatter Graph

The solid lines in figure 1 show the original CSA divisions while the broken lines show the divisions of the simplified categories. The WA\_RATIO label refers to the Wholist-Analytic dimension and the VI\_RATIO label refers to the Verbal-Imagery dimension.

**Table 2**  
Cognitive Style of Subjects (original categories)

**Table 3**  
Simplified Cognitive Style of Subjects

As shown table 2 and table 3, less than half of the subjects were within the middle section of either dimension, however, more than half the subjects were in the middle section of at least one dimension. Figure 1 identifies one extreme Analytic subject along the Wholist-Analytic dimension.

## Results Of All Questions

The descriptive results for each cognitive style group (the mean performance ratios of each group, the number of subjects in each group, and the standard deviation of each group) and the statistical significance of the descriptive results are presented for all the different question types in separate tables. The descriptive result tables are shown in two parts. The first part shows the results viewed from the two dimensions (the Wholist-Analytic dimension, and Verbal-Imagery dimension) and the mean results of all fifty subjects ('Total'). The second part shows the results of the four quadrants formed by combining the two dimensions (Wholist/Verbalisers are labelled WV, Wholist/Imagers are labelled WI, Analytic/Verbalisers are labelled AV and Analytic/Imagers are labelled AI). **The figures in brackets are the relative rankings of the cognitive style groups**. Wholists are ranked against Analytics, Verbalisers are ranked against Imagers, and the cognitive style quadrants are ranked against each other.

**Table 4**  
All Audio Questions Performance Ratios

Table 4 shows the subjects' performance ratios for all the questions asked in the audio test. While examining each cognitive style dimension individually Wholists performance ratio (9.50) is higher than the Analytics (9.15) and the Verbalisers performance ratio (9.32) is higher than the Imagers (9.30). This suggests that Wholists performed better than Analytics and Verbalisers performed better than Imagers. However, when dividing the sample into the four quadrants that are formed by combining the two dimensions Verbalisers do not always perform better than Imagers. While Verbalisers perform better than Imagers overall the Wholist/Imagers perform better than the Wholist/Verbalisers. This result shows that it is not always possible to draw conclusions by analysing the results of one dimension in isolation as the influence of the other dimension can also effect the results.

When comparing the descriptive results it is possible to rank the different cognitive style groups by their performance, however, the mean results do not establish whether the results of one group are significantly different from the results of another group. Therefore the Analysis of Variance (ANOVA) technique was used to calculate the between-subjects effects in order to determine whether the cognitive style has a significant influence on the results. For each cognitive style dimension where there is a 0.05 or less level of significance in the between-subjects effects then it is highly unlikely that the relative ranking of the cognitive style groups was caused by random chance.

The between-subjects effects test the significance of the differences between the results of each cognitive style dimension. For the Wholist-Analytic dimension ('WA') and the Verbal-Imagery dimension ('VI') the table shows the degrees of freedom ('df'), the F statistic, the level of significance ('Sig'), Eta squared and the observed power of the test. The Eta squared value measures the magnitude of effect, or the degree to which the variance in the results of subjects can be attributed to the cognitive style group to which they belong (Howell 1995). The results of the Observed Power test whether there is possibility of a Type II error of concluding that there is no significant difference between the samples where there is one. This value is calculated as 1 minus the probability of error, so the closer to 1 the value the less the probability of an error (Howell 1995).

**Table 5**  
Between-subjects Effects For All Audio Questions Performance Ratios

The between-subjects effects for all audio questions performance ratios (Table 5) were examined to test the significance of the differences between the results of each cognitive style dimension. The significance figures for both the both dimensions was above the 0.05 significance level and so while the descriptive results indicate that Wholists performed better than Analytics and Verbalisers performed better than Imagers there was no significant differences between the distribution of results for the either half of both dimensions. The Eta squared figures for both dimensions were very low which indicates that variance in the results are not due to cognitive style of subjects. The results of the Observed Power, however, are also low which suggests that there is a strong possibility of a type II error.

The one-way ANOVA post-hoc Tukey multiple comparison test was performed in order to determine whether the results of each of the cognitive style quadrants were significantly different from each other.

**Table 6**  
All Audio Questions Tukey Multiple Comparisons By Cognitive Style

Table 6 shows the results of the Tukey calculations for all audio questions. Results of each cognitive style quadrant (I) are compared to the results of the other three quadrants (J) in turn. For the results of two cognitive style quadrants to be significantly different the significance value should be 0.05 or less. In the above table none of the cognitive style quadrants are significantly different from any of the others.

The results of each cognitive style can be ranked against each other using the descriptive results, but the ANOVA results suggest that the results could be due to random chance. However, the low observed power of the test stops us from concluding that cognitive style has no effect on the results. One factor influencing this result may have been the combined effect of averaging the results of the different tasks that are suited for different cognitive styles. More significant differences between the cognitive style groups may be detected by examining more the different types of task. The next section analyses the results split into the 'S' and 'NS' questions

## **Comparison Of 'Speech' And 'Non-Speech' Question Results**

The relative results of the 'S' (Table 7) and 'NS' (Table 8) questions show that performance of subjects using audio is not consistent for all the cognitive style groups as performance is dependant on the type of question. Also the type of question that subjects perform better at corresponds to what is expected for their cognitive style (i.e. Verbalisers perform better at 'S' questions and the Imagers perform better at the 'NS' questions).

**Table 7**  
'Speech' Questions Performance Ratios

**Table 8**  
Simplified Cognitive Style of Subjects

As discussed in section 3.1, all subjects performed better at the 'S' questions than the 'NS'. Wholists had better performance ratios than Analytics in both the 'S' and 'NS' questions despite the fact that Analytics scored more correct answers in both cases. The stronger influence on performance ratio appears

to have been the Verbal-Imagery dimension as in the 'S' questions (Table 7) subjects in the Verbaliser quadrants performed better than the subjects in the Imager quadrants regardless of their Wholist-Analytic classification. Conversely, in the 'NS' questions (Table 8) subjects in the Imager quadrants performed better than Verbalisers regardless of their Wholist-Analytic classification. Within both the Verbaliser and Imager quadrants the subjects in the Wholist quadrants (1<sup>st</sup> and 3<sup>rd</sup>) performed better than the subjects in the Analytic quadrants (2<sup>nd</sup> and 4<sup>th</sup>).

### **Table 9**

#### Average 'Speech' / 'Non-Speech' Questions Performance Ratios

Averaging the results of the 'S'/'NS' questions (Table 9) did not change the relative rankings of the cognitive style groups from the results of all the audio questions (Table 4). There were more 'S' questions than 'NS' questions in the complete test and so the average shifted the overall results (Table 4) towards the 'NS' results (Table 8).

The performance ratios for the average 'S' and 'NS' questions (Table 9) show that Wholists performed better than Analytics regardless of their Verbal-Imagery classification. The performance of the Verbalisers and Imagers in the quadrants was more varied as the Wholist/Imagers performed best while the Analytic/Imagers performed worst.

### **Table 10**

#### Between-subjects Effects For 'Speech' Questions Performance Ratios

### **Table 11**

#### Between-subjects Effects For 'Non-Speech' Questions Performance Ratios

### **Table 12**

#### Between-subjects Effects For Average 'Speech' / 'Non-Speech' Questions Performance Ratios

For the 'S' questions (Table 10) the 'NS' questions (Table 11) and the average 'S' / 'NS' questions (Table 12) there were no significant differences in the subjects' performance ratios between the different cognitive style groups. Similarly, the results of the Tukey multiple comparison tests failed to find significant differences between the results of the cognitive style quadrants. This would suggest that there are no significant differences between the performance of results of cognitive style groups. The Eta

squared figures were low and did not indicate that the variance of the results was due to the cognitive style group that subjects belonged to. The low observed power values (less than 0.12 in all three of the above tables), however, suggest that there is a strong possibility that a type II error may have occurred in this case.

**Table 13**  
Repeated Measures Between-subjects Effects Comparing 'Speech' with 'Non-Speech' Questions  
Performance Ratios

In addition to the ANOVA calculations the individual types of task a General Linear Model Repeated Measures calculation was performed comparing the results of 'S' questions with the 'NS' questions (Table 13). A significant result of 0.05 or less occurs where the performance of subjects' in both tasks is significantly influenced by cognitive style, and the effect is significantly different in either task (e.g. where Verbalisers perform significantly better than Imagers in one task, and significantly worse in the other). Table 13 does not reveal that the way cognitive style effected subjects performance in the 'S' questions was associated with a corresponding effect in the 'NS' questions. This result was not unexpected as the between-subjects effects for the 'S' and 'NS' questions did not reveal any significant differences in performance for either types of task.

The results of 'S', 'NS' and average 'S/NS' questions like the all questions section do not show any significant differences between the results of the cognitive style groups. Again, however, it is not possible to declare that there were no significant differences between the results of the cognitive style groups because of the low observed power of the test. While examining the results for the totals of each types of task no significant differences between the performance of the cognitive style groups were detected. In the next two sections the results of the individual tasks are examined to discover if any significant results were hidden when aggregating the results.

### **'Speech' Questions**

The 'S' questions contained comparison questions, comprehension questions and mathematics questions. It was expected that Verbalisers would perform better at the questions that compared concepts and Imagers would perform better at the questions that compared objects (Riding 1991; Riding 1996).

**Table 14**  
Comparison Of Concepts Questions Performance Ratios

**Table 15**  
Comparison Of Objects Questions Performance Ratios

The relative performance of the Verbalisers and Imagers show a clear preference for the type of questions designed to be suited for their cognitive style. In the *comparison of concepts questions* the Verbalisers scored more than the Imagers, took less time than the Imagers and had higher performance ratio than the Imager subjects (Table 14). Similarly, in the *comparison of objects questions* the Imagers scored more than the Verbalisers, took less time than Verbalisers and had higher performance ratio than Verbalisers (Table 15).

**Table 16**  
Between-subjects Effects For Comparison Of Concepts Questions Performance Ratios

**Table 17**  
Between-subjects Effects For Comparison Of Objects Questions Performance Ratios

The analysis of the between-subjects effects did not find any significant differences between the results of the cognitive style groups for either the *comparison of concepts questions* (Table 16) or the *comparison of objects questions* (Table 17). Similarly, the results of the Tukey multiple comparison tests failed to find significant differences between the results of the cognitive style quadrants. The Eta squared figures for the Verbal-Imagery dimension in the *comparison of objects questions* were low and did not suggest that the variance in the results was due to cognitive style. The observed power figures were low in both tables which also indicates that there is a possibility of a type II error.

Although no significant effects were detected for the results of the performance ratios there was a marginally significant difference in the number of correct answers in the *comparison of concepts questions* for the Wholist-Analytic dimension ( $F = 2.80$ ,  $df = 1$ ,  $p = 0.10$ ). The Eta Squared value was low (0.06) but the observed power level was fairly low (0.37). Analytics scored a Mean of 74.07% compared to the Wholist Mean score of 70.65%. This effect showed that subjects' Wholist-Analytic classification was influencing their performance although there was no significant effect on the Verbal-Imagery dimension where an effect was expected to be seen.

**Table 18**

Repeated Measures Between-subjects Effects Comparing Concepts with Objects Questions Performance Ratios

The General Linear Model repeated measures calculation was performed to compare the results of the *comparison of concepts questions* and the results of the *comparison of objects questions* (Table 18). Although the descriptive results showed that Verbalisers performed better than Imagers at the *comparison of concepts questions* (Table 14) and Imagers performed better at the *comparison of objects questions* (Table 15) this was not a significant effect (Table 18).

**Table 19**

Procedural Questions Performance Ratios

In the *procedural questions* there was a wider range of scores than the overall audio questions, and so the scores influenced the performance ratio more than the duration figures (Table 19). The Analyticians scored more than Wholists and although they took longer than the Wholists they had a better performance ratio. Verbalisers scored more than the Imagers, took less time to answer the questions and had a better performance ratio than the Imagers. This result was as expected because words were being used in the questions (Riding 1996). Subjects in the Verbaliser quadrants also scored more than the Imager quadrants, took less time than the Imager quadrants and had better performance ratios than the Imager quadrants.

**Table 20**

Between-subjects Effects For Procedural Questions Performance Ratios

The between-subjects effects analysis for the *procedural questions* produced a significant result for the Verbal-Imagery dimension ( $F = 5.07$ ,  $df = 1$ ,  $p = 0.03$ ). The Eta squared value and the observed power value were also relatively high compared to all the other results in this experiment (Eta squared = 0.10, Observed Power = 0.60) which supports this finding. Verbalisers performed better than Imagers (Table 19) and their results were significantly different from each other (Table 20). The results of the Tukey multiple comparison tests, however, failed to find significant differences between the results of the cognitive style quadrants.

**Table 21**  
Mathematics Questions Performance Ratios

The results of the *Mathematics questions* (Table 21) do not show that either cognitive style dimension is more dominant than the other in influencing subjects' results. Scores for all the cognitive style categories were high and so the duration figures had more influence on determining the performance ratios. The Wholists scored more than Analytics, took less time to complete the questions and had a higher performance ratio. The Imagers scored more, but the Verbalisers took less time and had the higher performance ratio.

**Table 22**  
Between-subjects Effects For Mathematics Questions Performance Ratios

The analysis of the between-subjects effects also does not find any significant difference between the results of the different cognitive style groups (Table 22).

In all type of 'S' questions Analytics scored more correct answers than Wholists (except for *mathematics questions* in which the score of both Wholists and Analytics was high), but took longer and had better performance ratios (except the *mathematics questions* and the *comparison of objects questions*). In all 'S' questions Verbalisers were faster than Imagers and the performance ratios were better, except for the *comparison of objects questions*.

Not all the results of the individual types of 'S' questions were significantly different, but there were indications that subjects' cognitive style was influencing performance. Verbalisers' performance ratios in the *procedural questions* were significantly higher than the Imagers, and Analytics scored (marginal) significantly more correct answers than Wholists in the *comparison of concepts questions*. In addition it is not possible to declare that there were no significant differences between the results of the cognitive style groups in the other 'S' tasks because of the low observed power of the test. Significant differences between the performance ratios of the cognitive style groups were not identified in the *comparison of concepts questions*, the *comparison of objects questions* and the *mathematics questions* where performance was high for all subjects (see Section 3.1). There was a significant difference between the performance ratios of the cognitive style groups in the *procedural questions* where the performance of all subjects was not so uniformly high and there was more variation in the performance of subjects. In the

next section the results of the individual 'NS' tasks are examined where overall performance was relatively lower and there may be more possibility of detecting significant effects.

## 'Non-Speech' Questions

**Table 23**

Identification of 1 Instrument Questions Performance Ratios

**Table 24**

Identification of 2 Instruments Questions Performance Ratios

In both the identification of *one* (Table 23) and *two* (Table 24) *musical instruments questions* the Analytics have higher performance ratios than Wholists and Imagers have higher performance ratios than Verbalisers. Analytics scored more than the Wholists, and although the Wholists took less time the Analytics had higher performance ratio. Similarly the Imagers scored more than the Verbalisers, and although the Verbalisers took less time the Imagers had higher Performance Ratio.

For the identification of *two instruments questions* the scores were quite low for all subjects and there was a wide variation throughout the cognitive style categories and so score more influence than the duration figures in determining the performance ratios. In both the *identification of one* and *two instruments questions* neither dimension had more influence over the performance of subjects.

**Table 25**

Between-subjects Effects For Identification of 1 Instrument Questions Performance Ratios

**Table 26**

Between-subjects Effects For Identification of 2 Instruments Questions Performance Ratios

For the recognition of *one instrument questions* the analysis of the between-subjects effects did not find any significant differences between the distribution of the subjects performance ratios in the different cognitive style groups in the significance or Eta squared figures (Table 25). The observed power figures for the *identification of one instrument questions* were very low for both dimensions indicating there is a possibility of a type II error.

There was a marginally significant result in the *identification of two instrument questions* (Table 26) in the Verbal-Imagery dimension ( $F = 3.62$ ,  $df = 1$ ,  $p = 0.06$ ). The Eta squared value of 0.07 was slightly

higher than most of the results in this experiment but is still low. The observed power value of 0.46 shows there is a probability of error is just under half. The effect of cognitive style in *the identification of two instrument questions* is significant for the Verbal-Imagery dimension when looking at the score ( $F = 4.86$ ,  $df = 1$ ,  $p = 0.03$ ,  $\text{Eta Squared} = 0.1$ ,  $\text{Observed Power} = 0.58$ ).

The *identification of two instrument questions* also showed a marginally significant result between the cognitive style quadrants. Subjects' scores in the Wholist/Verbaliser quadrant (30.00%) were marginally significantly different from the subjects' scores in the Analytic/Imager quadrant (47.50%) (Mean Difference = -0.20, St. Error = 0.08,  $p = 0.07$ ).

**Table 27**  
Identification of Sound Effect Questions Performance Ratios

The scores for each cognitive style group for the *identification of sound effects questions* were high and relatively close to each other (between 70% for the Analytic/Verbalisers and 74.38% for the Wholist/Imagers) and so the duration figures had more influence on the performance ratios (Table 27). The Analytics scored more than the Wholists, but the Wholist took less time than the Analytics and had the higher performance ratio. Similarly the Imagers scored more than the Verbalisers, but the Verbalisers took less time to answer the questions than the Imagers and had the higher performance ratio.

**Table 28**  
Between-subjects Effects For Identification of Sound Effect Questions Performance Ratios

The between-subjects effects for the *identification of sound effects questions* performance ratios (Table 28) do not show any significant differences between performance ratios of the cognitive style groups. The Eta squared figures for dimensions are low. The Observed Power for both dimensions suggests that there is a strong possibility of a type II.

In all the 'NS' type of questions Analytics scored more but took longer than Wholists. Similarly in all the 'NS' type of questions Imagers scored more but took longer than Verbalisers. For the *identification of one and two instruments questions* Analytics perform better than Wholists and Imagers better than Verbalisers. The reverse is true only for the *identification of sound effect questions*.

There was a marginally significant difference between the results of Verbal-Imagery dimension cognitive style groups in the *identification of two instrument questions* with Imagers performing better

than Wholists. The Tukey multiple comparison tests also identified a marginally significant difference between the results of the Wholist/Verbaliser quadrant and the Analytic/Imager quadrant for scores in the *identification of two instrument questions*. For both the 'S' and 'NS' questions the significant results were detected in the types of task that all subjects performed relatively poorly as identified in Section 3.1.

In addition to analysing the results for each type of question all questions were assessed to determine which cognitive style each question is expected to be suitable for. The results of the analysis of the performance of subjects in questions that they were expected to perform relatively well and the questions they were not expected to perform relatively well are given in the next section.

## Comparison of Subjects' Performance By Cognitive Style

All questions used in the experiment were assessed in order to determine which cognitive group was expected to perform relatively better. Whether questions are suitable for either Wholists or Analytics is more dependent on the structure that information is presented, while whether questions are suitable for Verbalisers or Imagers depends on the types of words used in the question (Riding 1991; Riding 1996; Riding 1998). It was anticipated that the *procedural questions* would be suitable for Wholists, while the Analytics would perform better at the *mathematics questions*. In general it was anticipated that Verbalisers would prefer the 'S' questions and the Imagers would prefer the 'NS' questions, but there were additional complications due to the types of words that were used in the 'S' questions. It was anticipated that Wholists would perform better in the *comparison of concepts questions*, the *mathematics questions* and the *procedural questions*, but not the *comparison of objects questions*. It was anticipated that Imagers would perform better at all the 'NS' type of questions and the *comparison of objects questions*.

The results of the questions that were expected to be suitable for Wholists or Analytics did not produce the expected results. In the questions that were expected to be suitable for Wholists (the column labelled 'W Questions' in Table 29) it was the Analytics that unexpectedly performed better. The Wholist questions were also expected to be suitable for Verbalisers, and the Verbalisers did perform better than the Imagers. The ranking of the cognitive style quadrants appeared to be influenced more by the effect of the Verbal-Imagery dimension than the Wholist-Imagery dimension as the two Verbaliser quadrants were ranked higher than the two Imager quadrants. The influence of the Wholist-Analytic dimension on the

relative performance of the quadrants was less strong with the Analytic quadrants being ranked first and third.

**Table 29**  
Descriptive Results By Cognitive Styles Totals

In the questions that were expected to be suitable for Analytics (the column labelled 'A Questions' in Table 29) again the cognitive style that was not expected to perform best out performed the other as the Wholists performed better than the Analytics. However, the cognitive style quadrant that performed the best was one of the Analytic quadrants (Analytic/Verbalisers). The Analytic questions were also expected to be suitable for Verbalisers and the Verbalisers did perform better than the Imagers and the cognitive style quadrant that performed the best was one of the Verbaliser quadrants (Analytic/Verbalisers). Neither cognitive style dimension appeared to influence the ranking of the cognitive style quadrants more than the other as the Analytic quadrants were ranked first and fourth while the Verbaliser quadrants were ranked first and third.

The results of the questions that were expected to be suitable for Verbalisers and Imagers did produce results that were expected. Verbalisers performed better than Imagers at the questions that were expected to be suitable for Verbalisers (the column labelled 'V Questions' in Table 29) and Imagers performed better than Verbalisers in the questions that were expected to be suitable for Imagers (the column labelled 'I Questions' in Table 29). These results were supported by the relative rankings of the cognitive style quadrants. The Verbaliser quadrants were ranked first and second for the questions that were expected to be suitable for Verbalisers and the Imagers quadrants were ranked first and second for the questions that were expected to be suitable for Imagers.

**Table 30**  
Between-Subjects Effects For All Audio Tasks By Cognitive Style Totals

Although the Wholists and Analytics did not perform as expected in the questions that were expected to be suitable for Wholists there was a marginally significant interaction between the performance of subjects and whether they were Wholist or Analytic ( $df = 1, F = 3.64, p = 0.06$ ) as shown in Table 30. There was, however, more of an effect depending on whether the subject was Verbaliser or Imager ( $df = 1, F = 6.51, p = 0.01$ ). This significant effect is supported by the relatively high Eta squared

value (0.12) and relatively high observed power value (0.70). Using the Tukey calculation the performance of subjects in the Analytic/Verbaliser quadrant, that was ranked first, was found to be significantly different from the performance of subjects in the Wholist/Imager quadrant, that was ranked fourth ( $p = 0.02$ ).

The descriptive results of the questions that were expected to be suitable for Analytics did not show a strong influence of any cognitive style dimension, and the results of the between-subjects effects support this conclusion as there was no significant effect of either cognitive style dimension, the Eta squared value was low and the observed power was low.

As expected Verbalisers did perform significantly better than Imagers at the questions that were expected to be suitable for Verbalisers ( $df = 1, F = 5.72, p = 0.02$ ). This significant effect is supported by the relatively high Eta squared value (0.11) and relatively high observed power value (0.65). The Tukey calculation found that the performance of subjects in the Wholist/Imager quadrant, that was ranked first, was significantly better than the performance of subjects in the Analytic/Verbaliser quadrant, that was ranked fourth ( $p = 0.05$ ).

Similarly, Imagers performed (marginally) significantly better than Verbalisers at the questions that were expected to be suitable for Imagers ( $df = 1, F = 3.53, p = 0.07$ ), although the Eta squared value is lower than when a significant effect was identified (0.07) and the observed power value is not as high (0.45).

**Table 31**  
Repeated Measures Tests By Cognitive Style Totals

There was a marginal significant result detected by comparing how subjects performed in the questions that were expected to suit Wholists with the questions that were expected to suit Analytics (Table 31). The performance of subjects in the Wholist and Analytic questions marginally interacted with whether subjects were Verbalisers or Imagers ( $df = 1, F = 2.99, p = 0.09$ ).

While the analysis of the performance of subjects in questions that were expected to suit different cognitive styles did not prove that the performance of subjects in all cognitive style groups could be predicted in all cases there were indications that the cognitive style classification of subjects was influencing performance. More significant or marginally significant interactions were detected by analysing the performance in tasks without reference to the cognitive style that questions were suited for.

The next section discusses in more detail some of the issues raised by the results of the results of this experiment.

## DISCUSSION

A number of issues that arose from studying the results of the experiment shall be discussed in this section, including:

1. Identifying which tasks that subjects perform well.
2. Defining an average measure for audio performance
3. Detecting significant differences in the results of the different cognitive style groups
4. Whether it is appropriate to adapt the audio aspect of multimedia computer systems to suit the cognitive style of users
5. Identification of areas that require further study

The first issue is the identification of the tasks that are performed better by all subjects. In section 3.1 the performance of subjects in each type of task were ranked against each other. While it can be concluded that subjects perform better at certain tasks it is not possible to claim that it is appropriate to use audio for the tasks that all subjects performed well and it is not appropriate to use audio for the tasks that were not performed well. Some of the tasks in which subjects did not perform well such as the identification of instruments or sound effects are the types of task that cannot be performed without using audio. In order to determine which tasks are suited to the use of audio it is necessary to compare performance of the tasks using audio against performance of the same tasks using other media.

While, in general, subjects performed better at the tasks that used 'S' than the tasks that used 'NS' not all the 'S' type tasks were ranked above the 'NS' type tasks. The performance of subjects in the *identification of sound effects questions* was ranked above the performance of subjects in the *mathematics questions* (while not being significantly different), and above the performance in the *procedural questions*.

In addition to whether the task uses 'S' or 'NS' audio another important factor that influenced the relative ranking of the subjects' performance in the different tasks was the level of the demands placed on the subjects in each type of task. The types of task that subjects performed best involved simple

recognition and matching, while the tasks that subjects did not perform as well were progressively more complex. The *comparison questions*, ranked first and second, involved recognition of two words and required 'yes' or 'no' answers. The *identification of sound effects questions*, ranked equal third, involves subjects matching a sound with one of a number of possible pictorial representations of the source of the sound. The *mathematics questions*, also ranked equal third, placed more demands on subjects by asking them to perform simple calculations. The types of task that were ranked lower placed more complex demands on the subjects. In the *identification of single musical instruments questions*, ranked equal fourth, subjects had to recall similar sounding instruments and distinguish between them. For the *procedural questions*, ranked equal fourth, subjects had to recall the procedural instructions in the correct order and make calculations. The task that was performed least well by all subjects was the *identification of two instrument questions*. This task was more complex than the *identification of one instrument questions* as there was an additional distraction of instruments being playing simultaneously.

The relative performance of subjects between the tasks appears to be due to a combination of whether the task contains 'S' or 'NS' questions and the level of cognitive demands of each type of task. Subjects perform better where the level of cognitive demands are low, and where the level of cognitive demands are the same between 'S' and 'NS' types of questions then subjects will perform better at the type of questions that use speech.

There were, however, other factors that may have had an influence of the results of the tasks. The experiment was conducted using a multimedia computer system. Audio was used to present information to subjects and asks the questions, but subjects entered their answers by selecting an appropriate piece of text or image. One of the assumptions of the experiment was that subjects' performance in answering the questions would not be affected by their ability to read the text or distinguish between images that represented the possible multiple-choice answers. This, however, may not have been the case. In the tasks that subjects performed well most subjects scored a high number of correct answers and so the time it took for subjects to answer the questions was important to distinguish between the performance of subjects. Subjects who were able to process the text or image information that was required to enter the answer the questions faster than other subjects had a performance advantage that was not related to their ability to process the audio information.

The second issue involved the definition of an average performance measure for audio. An average performance for the two types of questions was derived. The overall audio performance measure of all

questions is not representative as a performance indicator for all subjects as the experiment was not designed to have equal numbers of all types of questions. Table 1 shows the wide variation of performance throughout all the tasks and the tests of between-subjects and showed that subjects performance in most of the tasks were significantly different. The distinction between the 'S' and 'NS' questions was chosen as a more representative indicator of average performance using audio. Although the relative ranking of the cognitive style groups is not different for all the audio questions (Table 4) and the average of the 'S' and 'NS' questions (Table 9) the two distributions of subjects results are significantly different from each other. The average figures indicate how subjects perform with an equal mix of 'S' and 'NS' audio information, however in a typical multimedia presentation the two types are often used in different ways. 'S' words are often used to provide information, while 'NS' sounds such as background music and sound effects are often used to create a certain mood or to highlight pieces of information that is being presented as speech or visually. Therefore, the performance figures for individual 'S' and 'NS' questions may still be used when weighted to reflect the audio mix, to deduce the performance of multimedia audio interfaces.

The third issue involves detecting significant differences between the results of the cognitive style groups. If there were significant differences then it would be appropriate to adapt the audio aspect of multimedia computer systems to suit the cognitive style of users. This, however, was not always the case and significant differences were the exception and not the rule. In most cases there were no significant differences between the results of the cognitive style groups. Where the observed power figures are low, however, there is a possibility of a type II error that a significant result has not been detected where there is an effect.

The significant or marginally significant results for differences in the performance ratios for each of the cognitive style groups occurred in the Verbal-Imagery dimension for the *procedural questions* where sig. = 0.03 (Table 20) and the *identification of two instruments questions* where sig. = 0.06 (Table 25). There was, however, no significance for the other tasks, although the observed power was low indicating that there was a high probability of a type II error. In addition, the Tukey calculations found no significant differences between the mean performance ratio of the cognitive style quadrants for any of the tasks.

No significant difference in the performance of subjects between the cognitive style groups were found in the tasks where the performance of all subjects were high, although it was not possible to totally rule out the presence of an effect because of the low observed power. Significant and marginally

significant effects were observed where the task was more demanding. This tends to suggest that the experiment used tasks that subjects found too easy, and if the experiment was more demanding more significant results would have been detected.

The fourth issue of whether it is appropriate to adapt the audio aspect of multimedia computer systems to suit the cognitive style of users depends on whether significant differences between the results of the cognitive style groups were detected. Looking at the descriptive results it appears that the mean performance does change, however, the results of all cognitive style groups are similar and few of the results are significantly different by the between-subjects tests. The results of the cognitive style groups were more significantly different in the more complex tasks regardless of whether they contained speech or not. If the differences in performance between the subjects are only apparent in the more complex tasks then it may be appropriate to adapt the tasks in a multimedia computer system to suit the subjects cognitive style.

There were no significant differences in the results of the *comparison of the objects or concepts questions* where differences may be expected to be found. This may be due to the simplicity of the tasks and differences can only be detected in more complex tasks, or that the multimedia interface of the experiment was not suited to detecting subtle differences between the different cognitive style group (subjects answered the audio questions by clicking on images or text). If the use of multimedia is affecting the results then it may not be appropriate to adapt the audio aspect of a multimedia computer system to the user's cognitive style because the other media used in the system may negate any anticipated benefit.

There are indications that performance is affected by the subjects' cognitive style, but the differences in the performance are slight. Significant differences were found where the task was more difficult and when the results of a number of tasks were added together (when assessing which questions were suitable for which cognitive style). This suggests that any benefit in adapting a computer system to suit the user's cognitive style will only be beneficial for difficult tasks, or where the system will be used over a long enough period of time for small effects to accrue. Further research needs to be undertaken in order to establish the extent that performance of different cognitive style groups is be affected in more complex tasks and over longer periods of time.

The fifth issue relates to the questions that are raised by results of this experiment, and what are the areas that require further study. Other media, including text, images and video, should be examined with similar tasks in order to compare which media are suited for the different types of task. Each experiment

should be designed in a manner that does not allow other media to contaminate the results. For example subjects answers for the audio questions should be given in a consistent manner without asking subjects to switch between selecting text for some questions and selecting images for other questions.

## CONCLUSIONS

We have devised an audio test using ‘Speech’ and ‘Non-Speech’ questions. The ‘Speech’ questions include *comparison of concepts questions, comparison of objects questions, mathematics questions and procedural questions*. The ‘Non-Speech’ questions include *identification of one or two musical instrument questions and the identification of sound effects questions*.

Subjects’ performance in the experiment was measured by recording score and duration in multiple-choice tests. We have adopted a new ratio to measure the performance of users when carrying out audio tasks using multimedia interfaces. The ratio of *correct score over duration* has been adopted as the performance measure of users for each type of interface.

The relative performances of subjects in each of the tasks were ranked against each other. Out of nine different types of task the results of all task were significantly different from each other except for two cases. Subjects performed best where the level of cognitive demands placed on them were low and perform progressively worse in the tasks that place greater cognitive demands on them. Comparing tasks with similar levels of cognitive demands subjects will perform better at the type of questions that use ‘Speech’ rather than ‘Non-Speech’.

Cognitive style was not found to be a significant influence the performance of subjects when looking at the overall performance in all questions, or when looking at the results of the overall results of the ‘Speech’ and ‘Non-Speech’ tasks. Indications that cognitive style was having a significant influence on the performance of subjects started to be detected when examining the results of the individual tasks. The significant results were not detected in the tasks in the relatively simple tasks which all subjects performed uniformly well but were detected in the more complex tasks that all subjects performed not so well and there was more variation in the performance of subjects. More significant results were detected when categorising the questions in terms of which cognitive style they were suited for. In particular the Verbal-Imagery dimension was found to influence performance. Where no significant difference were found in the

results of cognitive style groups it was possible that there was an effect but it was not detected because the low observed power of the test.

Developers of multimedia computer systems can use the results of the experiment. Performance of all subjects regardless of cognitive style can be predicted for each type of task. The relative performance of subjects in different cognitive style groups cannot be predicted for the relatively simple tasks but the relative performance of Verbalisers and Imagers can be predicted for more complex tasks. This result suggests that the benefit to users of adapting the audio aspect of multimedia computer systems increases with the complexity of the task.

## References

- Beaumont, I. H. (1994). User Modelling in the Interactive Anatomy Tutoring System ANATOM-TUTOR. *User Modeling and User-Adapted Interaction* 4(1), 21-45.
- Benyon, D. and D. Murray (1993). Applying user modelling to human-computer interaction design. *Artificial Intelligence Review* 7, 199-225.
- Bocker, H.-D., H. Hohl, et al. (1990). HYPADAPTER - Individualising Hypertext human-computer interaction. *Human-Computer Interaction - INTERACT 90*. Ed. Diaper, D., Elsevier Science Publishers B. V. (North Holland) 931-936.
- Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *User-Adapted Interaction* 6, 87-129.
- Elson-Cook, M. (1993). Student modelling in intelligent tutoring systems. *Artificial Intelligence Review* 7, 221-240.
- Goldstein, E. B. (1989). *Sensation and Perception*. Brooks/Cole.
- Gross, R. D. (1992). *Psychology The science of Mind and Behaviour*. Hodder & Stoughton.
- Handel, S. (1993). *Listening: An Introduction to the Perception of Auditory Events*. A Bradford Book, The MIT Press.
- Hohl, H., H.-D. Bocker, et al. (1996). Hypadapter: An Adaptive Hypertext System for Exploratory Learning and Programming. *User Modelling and User-Adapted Interaction* 6, 131-156.
- Howell, D. C. (1995). *Fundamental Statistics For The Behavioural Sciences*, 3<sup>rd</sup> Edition, Duxbury Press.
- John, D., J. Jerrams-Smith, et al. (1998). Telecare Companion: An Adaptive Telecare System. *SCSDSP 98*, Sheffield Hallam University Press.
- Kolb, D. A. and R. Fry (1975). Towards an Applied Theory of Experiential Learning. In *Theories of Group Processes*. Ed. Cooper, C. L., Pitman Press: 33 - 57.
- Riding, R. (1991). *Cognitive Styles Analysis Users' Manual*. Learning & Training Technology.
- Riding, R. and I. Cheema (1991). Cognitive Styles - an overview and integration. *Educational Psychology* 11(3 and 4), 193 - 215.

Riding, R. and E. Sadler-Smith (1992). Type of Instructional Material, Cognitive Style and Performance. *Educational Studies* 18(3), 323-340.

Riding, R. (1994). *Personal Style Awareness and Personal Development*. Birmingham, Learning & Training Technology.

Riding, R. J. (1996). *Learning Styles and Technology-Based Training*. Sheffield, Department For Education and Employment.

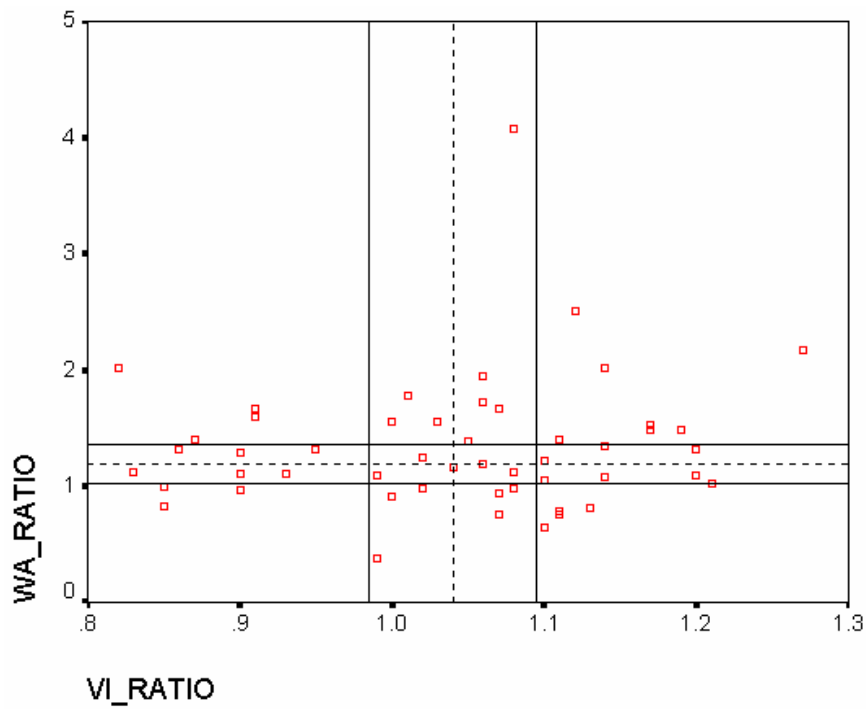
Riding, R. J., Glass et al. (1997). Cognitive Style and Individual Differences in EEG Alpha During Information Processing. *Educational Psychology* 17(1 and 2), 219 – 234

Riding, R. J. and A. Staley (1998). Self-perception as Learner, Cognitive Style and Business Studies Students' Course Performance. *Assessment & Evaluation in Higher Education* 23(1), 43 - 58.

Riding, J. (1998). *Cognitive Styles Analysis: Research Applications*. Birmingham, Learning & Training Technology.

Tennant, M (1988). *Psychology & Adult Learning*. Routledge

Acknowledgements: We would like to thank BT Labs, Ipswich, UK, for kindly sponsoring this project.



**Figure 1.** Cognitive Style Of Subjects Scatter Graph

**Table 1**

Mean Performance Ratios For All Audio Experiment Tasks

Task	Mean	N	Std. Dev.
All Questions	9.31	50	1.57
S	11.12	50	1.86
NS	7.54	50	1.97
Average S/NS	9.20	50	1.58
Compare Concepts	12.09	50	2.29
Compare Objects	14.50	50	2.72
Procedural	7.95	50	2.17
Maths	10.00	50	2.84
Identify 1 Instrument	8.78	50	3.57
Identify 2 Instruments	3.22	50	1.95
Identify Sound Effects	10.78	50	3.19

**Table 2**

Cognitive Style of Subjects (original categories)

Cognitive Style	Verbaliser	Bimodal	Imager	Total
Wholist	3	6	5	14
Inter-mediate	6	5	6	17
Analytic	4	8	7	19
Total	13	18	19	50

**Table 3**  
Simplified Cognitive Style of Subjects

Cognitive Style	Verbaliser	Imager	Total
Wholist	10	13	23
Analytic	11	16	27
Total	21	29	50

**Table 4**  
All Audio Questions Performance Ratios

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev.
Wholist	9.50 (1)	23	1.61
Analytic	9.15 (2)	27	1.55
Verbaliser	9.32 (1)	21	1.54
Imager	9.30 (2)	29	1.62
Total	9.31	50	1.57
WV	9.49 (2)	10	1.68
WI	9.51 (1)	13	1.62
AV	9.17 (3)	11	1.46
AI	9.13 (4)	16	1.65

**Table 5**  
Between-subjects Effects For All Audio Questions Performance Ratios

Cognitive Style Dimension	df	F	Sig.	Eta Squared	Observed Power
WA	1	0.57	0.46	0.01	0.11
VI	1	0.00	0.99	0.00	0.05

**Table 6**  
All Audio Questions Tukey Multiple Comparisons By Cognitive Style

Cognitive Style Quadrants		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I)	(J)				Lower Bound	Upper Bound
WV	WI	-0.03	0.68	1.00	-1.83	1.77
	AV	0.31	0.70	0.97	-1.56	2.18
	AI	0.36	0.65	0.95	-1.37	2.08
WI	WV	0.03	0.68	1.00	-1.77	1.83
	AV	0.34	0.66	0.95	-1.42	2.10
	AI	0.38	0.60	0.92	-1.22	1.99
AV	WV	-0.31	0.70	0.97	-2.18	1.56
	WI	-0.34	0.66	0.95	-2.10	1.42
	AI	0.04	0.63	1.00	-1.63	1.72
AI	WV	-0.36	0.65	0.95	-2.08	1.37
	WI	-0.38	0.60	0.92	-1.99	1.22
	AV	-0.04	0.63	1.00	-1.72	1.63

**Table 7**  
‘Speech’ Questions Performance Ratios

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev.
Wholist	11.23 (1)	23	1.88
Analytic	11.03 (2)	27	1.87
Verbaliser	11.32 (1)	21	2.18
Imager	10.97 (2)	29	1.62
Total	11.12	50	1.86
WV	11.53 (1)	10	2.46
WI	11.00 (3)	13	1.35
AV	11.13 (2)	11	1.99
AI	10.96 (4)	16	1.85

**Table 8**  
Simplified Cognitive Style of Subjects

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev
Wholist	7.74 (1)	23	2.11
Analytic	7.36 (2)	27	1.86
Verbaliser	7.28 (2)	21	1.65
Imager	7.72 (1)	29	2.18
Total	7.54	50	1.97
WV	7.31 (3)	10	1.73
WI	8.07 (1)	13	2.37
AV	7.25 (4)	11	1.66
AI	7.45 (2)	16	2.04

**Table 9**  
Average ‘Speech’ / ‘Non-Speech’ Questions Performance Ratios

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev.
Wholist	9.39 (1)	23	1.62
Analytic	9.03 (2)	27	1.55
Verbaliser	9.19 (2)	21	1.52
Imager	9.20 (1)	29	1.64
Total	9.20	50	1.58
WV	9.35 (2)	10	1.66
WI	9.42 (1)	13	1.65
AV	9.05 (3)	11	1.46
AI	9.02 (4)	16	1.66

**Table 10**  
Between-subjects Effects For ‘Speech’ Questions Performance Ratios

Cognitive Style Dimension	df	F	Sig.	Eta Squared	Observed Power
WA	1	0.16	0.69	0.00	0.07
VI	1	0.43	0.52	0.01	0.10

**Table 11**

Between-subjects Effects For 'Non-Speech' Questions Performance Ratios

Cognitive Style Dimension	df	F	Sig.	Eta Squared	Observed Power
WA	1	0.35	0.56	0.01	0.09
VI	1	0.69	0.41	0.02	0.13

**Table 12**

Between-subjects Effects For Average 'Speech' / 'Non-Speech' Questions Performance Ratios

Cognitive Style Dimension	df	F	Sig.	Eta Squared	Observed Power
WA	1	0.57	0.45	0.01	0.12
VI	1	0.00	0.97	0.00	0.05

**Table 13**

Repeated Measures Between-subjects Effects Comparing 'Speech' with 'Non-Speech' Questions Performance Ratios

Cognitive Style Dimension	df	F	Sig.
WA	1	0.39	0.54
VI	1	0.02	0.89

**Table 14**

Comparison Of Concepts Questions Performance Ratios

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev.
Wholist	11.87 (2)	23	2.51
Analytic	12.28 (1)	27	2.12
Verbaliser	12.33 (1)	21	1.96
Imager	11.92 (2)	29	2.53
Total	12.10	50	2.29
WV	12.04 (3)	10	2.38
WI	11.74 (4)	13	2.70
AV	12.60 (1)	11	1.55
AI	12.06 (2)	16	2.47

**Table 15**

Comparison Of Objects Questions Performance Ratios

Cognitive Style	Mean Perf. Ratio	Number in group	St. Dev.
Wholist	14.74 (1)	23	2.76
Analytic	14.30 (2)	27	2.72
Verbaliser	13.96 (2)	21	3.22
Imager	14.89 (1)	29	2.27
Total	14.50	50	2.72
WV	14.53 (3)	10	3.46
WI	14.90 (1)	13	2.23

<b>AV</b>	13.44 (4)	11	3.05
<b>AI</b>	14.88 (2)	16	2.38

**Table 16**

Between-subjects Effects For Comparison Of Concepts Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.43	0.52	0.01	0.10
<b>VI</b>	1	0.38	0.54	0.01	0.09

**Table 17**

Between-subjects Effects For Comparison Of Objects Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.49	0.49	0.01	0.11
<b>VI</b>	1	1.33	0.26	0.03	0.20

**Table 18**

Repeated Measures Between-subjects Effects Comparing Concepts with Objects Questions Performance Ratios

<b>Cognitive Style</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>
<b>WA</b>	1	0.01	0.93
<b>VI</b>	1	0.16	0.69

**Table 19**

Procedural Questions Performance Ratios

<b>Cognitive Style</b>	<b>Mean Perf. Ratio</b>	<b>Number in group</b>	<b>St. Dev</b>
<b>Wholist</b>	7.94 (2)	23	2.17
<b>Analytic</b>	7.96 (1)	27	2.22
<b>Verbaliser</b>	8.73 (1)	21	2.16
<b>Imager</b>	7.38 (2)	29	2.04
<b>Total</b>	7.95	50	2.17
<b>WV</b>	8.89 (1)	10	2.18
<b>WI</b>	7.20 (4)	13	1.93
<b>AV</b>	8.60 (2)	11	2.23
<b>AI</b>	7.53 (3)	16	2.17

**Table 20**

Between-subjects Effects For Procedural Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed</b>
----------------------------------	-----------	----------	-------------	--------------------	-----------------

					<b>Power</b>
<b>WA</b>	1	0.00	0.98	0.00	0.05
<b>VI</b>	1	5.07	0.03	0.10	0.60

**Table 21**  
Mathematics Questions Performance Ratios

<b>Cognitive Style</b>	<b>Mean Perf. Ratio</b>	<b>Number in group</b>	<b>St. Dev</b>
<b>Wholist</b>	10.13 (1)	23	2.66
<b>Analytic</b>	9.89 (2)	27	3.03
<b>Verbaliser</b>	10.35 (1)	21	2.94
<b>Imager</b>	9.74 (2)	29	2.78
<b>Total</b>	10.00	50	2.84
<b>WV</b>	10.07 (3)	10	3.49
<b>WI</b>	10.17 (2)	13	1.95
<b>AV</b>	10.60 (4)	11	2.49
<b>AI</b>	9.40 (2)	16	3.34

**Table 22**  
Between-subjects Effects For Mathematics Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.02	0.88	0.00	0.05
<b>VI</b>	1	0.44	0.51	0.01	0.10

**Table 23**  
Identification of 1 Instrument Questions Performance Ratios

<b>Cognitive Style</b>	<b>Mean Perf. Ratio</b>	<b>Number in group</b>	<b>St. Dev</b>
<b>Wholist</b>	8.59 (2)	23	3.41
<b>Analytic</b>	8.93 (1)	27	3.76
<b>Verbaliser</b>	8.56 (2)	21	3.78
<b>Imager</b>	8.93 (1)	29	3.47
<b>Total</b>	8.78	50	3.57
<b>WV</b>	7.78 (4)	10	3.53
<b>WI</b>	9.22 (2)	13	3.32
<b>AV</b>	9.28 (1)	11	4.02
<b>AI</b>	8.70 (3)	16	3.68

**Table 24**  
Identification of 2 Instruments Questions Performance Ratios

<b>Cognitive Style</b>	<b>Mean Perf. Ratio</b>	<b>Number in group</b>	<b>St. Dev</b>
<b>Wholist</b>	3.15 (2)	23	1.80
<b>Analytic</b>	3.28 (1)	27	2.09
<b>Verbaliser</b>	2.60 (2)	21	1.39
<b>Imager</b>	3.66 (1)	29	2.18

<b>Total</b>	3.22	50	1.95
<b>WV</b>	2.58 (4)	10	1.81
<b>WI</b>	3.58 (2)	13	1.74
<b>AV</b>	2.62 (3)	11	0.96
<b>AI</b>	3.73 (1)	16	2.54

**Table 25**

Between-subjects Effects For Identification of 1 Instrument Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.22	0.64	0.01	0.07
<b>VI</b>	1	0.17	0.68	0.00	0.07

**Table 26**

Between-subjects Effects For Identification of 2 Instruments Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.03	0.87	0.00	0.05
<b>VI</b>	1	3.62	0.06	0.07	0.46

**Table 27**

Identification of Sound Effect Questions Performance Ratios

<b>Cognitive Style</b>	<b>Mean Perf. Ratio</b>	<b>Number in group</b>	<b>St. Dev</b>
<b>Wholist</b>	11.15 (1)	23	3.85
<b>Analytic</b>	10.46 (2)	27	2.53
<b>Verbaliser</b>	10.91 (1)	21	3.24
<b>Imager</b>	10.68 (2)	29	3.20
<b>Total</b>	10.78	50	3.19
<b>WV</b>	12.12 (1)	10	3.83
<b>WI</b>	10.41 (3)	13	3.84
<b>AV</b>	9.81 (4)	11	2.22
<b>AI</b>	10.90 (2)	16	2.70

**Table 28**

Between-subjects Effects For Identification of Sound Effect Questions Performance Ratios

<b>Cognitive Style Dimension</b>	<b>df</b>	<b>F</b>	<b>Sig.</b>	<b>Eta Squared</b>	<b>Observed Power</b>
<b>WA</b>	1	0.99	0.33	0.02	0.16
<b>VI</b>	1	0.11	0.74	0.00	0.06

**Table 29**  
Descriptive Results By Cognitive Styles Totals

<b>Cognitive Style</b>		<b>W Questions</b>	<b>A Questions</b>	<b>V Questions</b>	<b>I Questions</b>
<b>Wholist</b>	<b>Mean</b>	8.50	10.39	9.41	11.49
	<b>Rank</b>	2.00	1.00	2.00	1.00
	<b>N</b>	23.00	23.00	23.00	23.00
	<b>Std. Dev.</b>	2.23	2.68	1.71	2.04
<b>Analytic</b>	<b>Mean</b>	9.71	10.01	10.11	11.47
	<b>Rank</b>	1.00	2.00	1.00	2.00
	<b>N</b>	27.00	27.00	27.00	27.00
	<b>Std. Dev.</b>	2.42	2.88	1.93	1.58
<b>Verbaliser</b>	<b>Mean</b>	10.07	10.40	10.48	10.92
	<b>Rank</b>	1.00	1.00	1.00	2.00
	<b>N</b>	21.00	21.00	21.00	21.00
	<b>Std. Dev.</b>	2.09	2.82	1.66	1.95
<b>Imager</b>	<b>Mean</b>	8.49	10.03	9.29	11.88
	<b>Rank</b>	2.00	2.00	2.00	1.00
	<b>N</b>	29.00	29.00	29.00	29.00
	<b>Std. Dev.</b>	2.41	2.77	1.85	1.57
<b>Total</b>	<b>Mean</b>	9.15	10.18	9.79	11.48
	<b>Rank</b>	2	1	2	1
	<b>N</b>	50.00	50.00	50.00	50.00
	<b>Std. Dev.</b>	2.39	2.77	1.85	1.79
<b>WV</b>	<b>Mean</b>	9.56	10.17	10.08	10.92
	<b>Rank</b>	2.00	3.00	2.00	3.00
	<b>N</b>	10.00	10.00	10.00	10.00
	<b>Std. Dev.</b>	1.79	3.28	1.57	2.23
<b>WI</b>	<b>Mean</b>	7.68	10.55	8.90	11.92
	<b>Rank</b>	4.00	2.00	4.00	1.00
	<b>N</b>	13.00	13.00	13.00	13.00
	<b>Std. Dev.</b>	2.25	2.24	1.69	1.85
<b>AV</b>	<b>Mean</b>	10.54	10.60	10.85	10.92
	<b>Rank</b>	1.00	1.00	1.00	4.00
	<b>N</b>	11.00	11.00	11.00	11.00
	<b>Std. Dev.</b>	2.31	2.49	1.72	1.77
<b>AI</b>	<b>Mean</b>	9.15	9.60	9.61	11.84
	<b>Rank</b>	3.00	4.00	3.00	2.00
	<b>N</b>	16.00	16.00	16.00	16.00
	<b>Std. Dev.</b>	2.41	3.14	1.96	1.37

**Table 30**

Between-Subjects Effects For All Audio Tasks By Cognitive Style Totals

Cognitive Style	Cognitive Style Dimension	df	F	Sig.	Eta Squared	Observed Power
Wholist	WA	1	3.64	0.06	0.07	0.46
	VI	1	6.51	0.01	0.12	0.70
Analytic	WA	1	0.10	0.75	0.00	0.06
	VI	1	0.15	0.70	0.00	0.07
Verbaliser	WA	1	2.11	0.15	0.04	0.30
	VI	1	5.72	0.02	0.11	0.65
Imager	WA	1	0.01	0.93	0.00	0.05
	VI	1	3.53	0.07	0.07	0.45

**Table 31**

Repeated Measures Tests By Cognitive Style Totals

Cognitive Style Comparisons	Cognitive Style Dimension	df	F	Sig.
W Questions * A Questions	WA	1	0.73	0.40
	VI	1	2.99	0.09
V Questions * I Questions	WA	1	0.80	0.38
	VI	1	0.11	0.74