An Experimental Evaluation of an Alternative to the Pivot Table for A*d Hoc* Access to OLAP Data

Peter O'Donnell
Nick Draper

Decision Support Systems Laboratory
Monash University
Melbourne, Australia
Email: peter.odonnell@infotech.monash.edu.au

Abstract

This paper examines the usability of the orthodox interface to access business intelligence data held in OLAP-based systems – the pivot table. An alternative to the pivot table is described, based on Erik Thomsen's simple diagramming technique for designing OLAP data structures. That interface is compared to the pivot table in a laboratory-based experiment. The results show that the alternative interface is a better interface to use for ad hoc access to OLAP data. The results for the subjects using the pivot table are very poor, with nearly a third of them being unable to successfully complete any of three simple analysis tasks. These results have important implications, as many systems and software tools are now based on the pivot table interface. The pivot table could be too difficult for most users to use, limiting the success of business intelligence systems based on OLAP technology.

Keywords

Business intelligence, OLAP, Pivot table, Laboratory experiment

1. INTRODUCTION

Online-analytic processing (OLAP) systems can provide managers and analysts with fast, intuitive and flexible access to important data. The aim of OLAP is to help users analyse and understand that data. The screens of an OLAP system show data that is often summarised, graphed and color-coded. Users can simply select a parameter, give it a new value, and wait for a moment while the screen refreshes to display new data based on the selected parameter. The fundamental software tool for *ad hoc* access to OLAP data has a flexible spreadsheet style interface; it is usually called a pivot table or a worksheet. Many vendors have developed tools based on the pivot table interface, which allows users to connect to a variety of OLAP data sources and then explore and manipulate them by dropping and dragging the 'dimensions' that define the data displayed.

When vendors demonstrate the functionality of these tools the audience is usually very impressed. Typically, the demonstration will act out a scenario, for example showing how a marketing manager is able to use OLAP to 'slice and dice' their way through the data from their corporate data warehouse to identify a poorly performing product or sales region. At the end of the demonstration, after the ooh's and ah's and the applause has stopped, sales are made. In short, OLAP and OLAP worksheets in particular 'demo' well. However, in reality the OLAP pivot table is a difficult tool to use. Only a small proportion of users in most OLAP installations will ever be regular and productive users of this interface. This presents the developers of OLAP systems with a significant problem as implementations that rely on a pivot table interface are not likely to succeed.

In this paper, an alternative interface for *ad hoc* access to OLAP data is presented and a prototype system based on that interface is discussed. The interface of this prototype system is based on a metaphor known as a multi-dimensional domain structure or more simply a Thomsen diagram, This metaphor was developed by Erik Thomsen to assist in the design of OLAP data structures (Thomsen, 1997). This paper examines the proposition that this metaphor can form the basis of an effective interface for *ad hoc* access to OLAP data.

The paper is structured as follows. The next section summarises the history of OLAP technology and discusses the rise to prominence of the pivot table. This section also discusses the nature of the pivot table interface and its weaknesses. This is followed by a description of an alternative interface based upon the multi-dimensional domain structure metaphor. The method of an experiment designed to compare this alternative metaphor to the pivot table is then described. The results are presented and discussed. The paper then concludes by discussing the implications of the results, which present a challenge to OLAP tool vendors to work to build better interfaces for OLAP systems.

2. OLAP AND THE PIVOT TABLE

The technology that is used to build OLAP systems is an important and increasingly widely used class of software. However, OLAP technology is not new. The first OLAP systems were developed in the early 1970s (Crandall, 1996). The technology was used to develop systems that provided senior managers with a briefing book summarising the performance of their business unit. These systems were often called executive information systems (EIS). They focused on displaying comparisons of key financial and operational data with target and budget figures. The user would "then select which line items to examine in more detail … this drill down facility allows him [sic] to explore the data which is contributing to exceptional variances and to seek out the factors requiring action" (Martin & Clarke, 1990).

The term OLAP was not coined to describe the technology used to build these systems until 1993 (Codd, Codd, & Salley, 1993). At that time the data warehousing movement, with similar aims to the EIS movement but a broader scope, began to gather momentum. For a time, the data warehousing movement swamped the OLAP 'industry'. Many data warehousing practitioners, consultants and vendors, who didn't understand OLAP, simply ignored it. This situation is changing. OLAP technology is complimentary to the aims (Kimball, 1996), methods, architectures and technologies used in data warehousing (Kimball, Reeves, Ross , & Thornwaite, 1999; Thomsen, 1997). Relational database vendors such as Microsoft, Oracle and IBM, in order to compete in the competitive and growing data warehousing and business intelligence markets, have included OLAP functionality within their core relational database products. The lowered ownership costs and integration of OLAP with relational database management systems (RDBMS) has broadened the OLAP market. It has also increased the level of understanding of OLAP among the database professionals who develop and maintain data warehouses. OLAP is now seen as an important component of the technology toolkit required to develop business intelligence systems.

This integration of OLAP and RDBMS technology has seen a major change in the nature and composition of the tools offered by OLAP vendors. Firms like Comshare, Pilot Software, Information Resources International, Planning Sciences and Holistic Systems once dominated the market. As more vendors entered the market, especially the large ones (Microsoft and Oracle in particular), these firms have either adapted and shifted focus, gone out of business, or been acquired by other firms. The change in the market hasn't been restricted to the names printed on the side of the software boxes or the lowering of the cost of software licenses; the nature of OLAP technology has changed.

It has been argued that in the past, two main obstacles prevented the widespread acceptance of OLAP technology: the volumes of data that they could handle were relatively small, and the data was stored in a proprietary format (Kimball, 1996). These issues have been addressed as OLAP technology has evolved. OLAP systems now have the capacity to handle large enough data volumes for most well designed applications. For example, Microsoft's Analysis Services product has been used to process an OLAP data structure based on a 1.2-terabyte data warehouse that contained 7.5 billion facts (EMC, Unisys, & Microsoft, 2002). OLAP is now also 'open'. The members of the OLAP Council – a consortium of OLAP vendors – were the first to develop an application-programming interface (API) to allow open access to OLAP data. This API is now redundant following the widespread adoption of the Microsoft developed and sponsored OLE DB for OLAP standard and the multi-dimensional expressions (MDX) language.

While the OLAP server market has standardised and consolidated as a result of the availability and adoption of standards like MDX, there has been an explosion in the number of OLAP clients available to access data stored on OLAP servers. These are available in a variety of forms including traditional workstation based clients, spreadsheet add-ins and web browser clients based on dynamic HTML. These clients are capable of linking to a variety of OLAP data sources, provided they comply with the OLE DB standard, and most offer a pivot table style interface to view and interact with the OLAP data. These pivot tables allow users to orient and filter the data they are viewing in any manner that they choose.

Pivot tables look much like a spreadsheet. It is not unusual for OLAP tool vendors to describe their product as "being like a spreadsheet on steroids" (Druker in McKendrick, 1998). They usually explain the use of a pivot table with an example that starts with a spreadsheet, explaining that its data display grid is essentially two-dimensional. This leads to a discussion of their pivot table showing how extra dimensions of data can be used to select and display data. The popular spreadsheet product – Microsoft Excel - has a pivot table function built into it. Spreadsheets are a familiar tool for most of the users (managers and analysts) of OLAP systems. As a result "many vendors of multidimensional tools have intentionally used spreadsheet products as interfaces"

(Thomsen, 1997). There are many Excel add-ins available that offer more advanced pivot table functionality than the basic pivot table built-in to the product.

A typical business intelligence system interface is shown in Figure 1. The system shown is the *business intelligence portal* from Microsoft, a web-based interface to OLAP data held in the Microsoft Analysis Services product (Microsoft, 2003). The screen shows, within a web browser, a pivot table showing product sales figures for a fictitious beverage company. Note that the data is displayed in a spreadsheet style grid. The data shown is from the OLAP database used by the XML for Analysis Council for interoperability testing (XMLA, 2003). In the figure a product dimension is shown in the rows section of the pivot table. One item of that dimension, in Figure 1, the 'Colas' product family, has been expanded to show dimension items at the SKU level. In the columns of the pivot table, financial measures named profit, sales and margin are shown. The data structure contains other dimensions: time, market geography, supplier and scenario. These are shown across the top of the pivot table as "paging" dimensions.

The user can, by interacting with the pivot table, change the data displayed. For example, the user can expand or collapse dimensional hierarchies – like the one expanded in the product family – by clicking the [+] and [-] boxes displayed in the row and column headings or by double clicking on individual row or column headings. Any of the paging dimensions can be 'dragged' by the user from the display bar at the top of the pivot table into the row and column area of the pivot table. Depending on where the drag ends, and whether or not a dimension has previously been set as a row or column dimension, this action might cause the dragged dimension to replace the current setting or to display its values nested within the existing dimension. Similarly, a dimension currently set as a row or column dimension can be dragged out and reset as a paging dimension. Each time the dimensions are rearranged by the user the data displayed is refreshed.

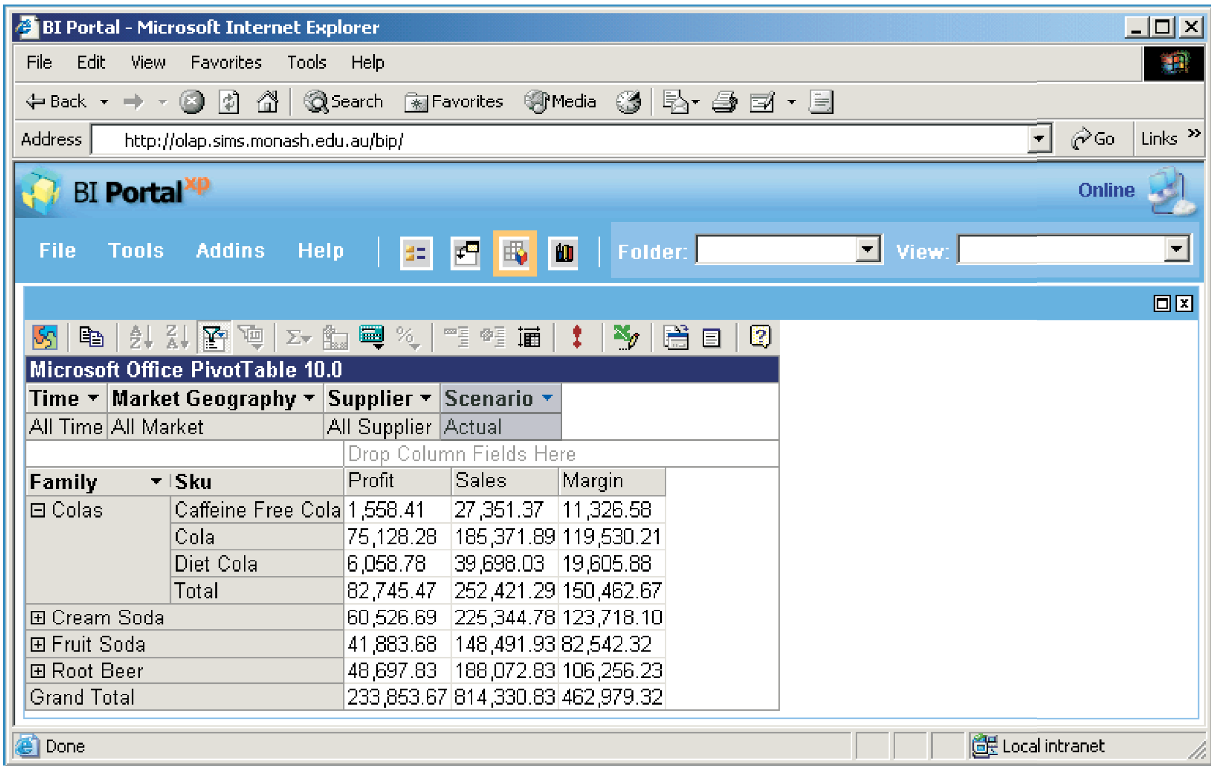

Figure 1: A typical business intelligence interface to OLAP data.

Very little administration work has to be performed to connect a pivot table based client, like the one shown in Figure 1, to an OLAP data source. Once the data source has been set up on an OLAP server and a user account created, a user connects their client tool - which in many cases might simply be a web browser - to the server, enters a username and a password, selects a data source and can immediately start to browse the data. There is very little difference in the basic functionality of the OLAP data browsers provided by dozens of different OLAP client tool vendors. In the last few years, the pivot table and the functionality it provides for data exploration has come to represent most people's understanding of what OLAP technology can provide users.

The software tool shown in Figure 2 was developed for the experiment described in this paper and was based upon the standard common object model (COM) pivot table provided by Microsoft. Note how similar it is to the web browser-based tool shown in Figure 1, both are based on the same standard COM component distributed similar tool with Microsoft's Office suite of end-user productivity tools. Many OLAP client applications have been developed using this software component.

However, despite its popularity and widespread distribution there are problems with the pivot table. In order to use a pivot table to search for data in an OLAP database the user has to do a lot of work. The default presentation of an OLAP structure in a pivot table – the starting point for analysis by users - might make sense for an accountant but it might be confusing for another user, for example a marketer. Different views of data look very similar and an incorrect setting, for example in a paging dimension used to filter the data, could easily be missed by a user. In a spreadsheet, users click in the data grid – the body of the spreadsheet – to manipulate data. In a pivot table, clicking within the data grid doesn't do anything. The functionality of the pivot table is invoked by interacting with the row and column headings. An experienced spreadsheet user might find this confusing as the row and column headings of a spreadsheet are merely place markers. Further, in order to understand the data being viewed the user must understand the complex multi-dimensional data structure they are manipulating. Most OLAP data structures are complex. For example the standard demonstration sales data structure shipped by Microsoft in their Analysis Services product and used in many books and tutorials on the use of OLAP has 11 dimensions. It is not unusual for production OLAP databases to have an even larger number of dimensions. The hierarchies within individual dimensions might also be a source of confusion or frustration for users. For example, a search for a specific product in a typical product dimension – where there could be tens of thousands of fields – might require a user to remember specific information about where a product has been located in a many leveled hierarchy of product families, groups, categories and lines in order to find the item they are looking for.

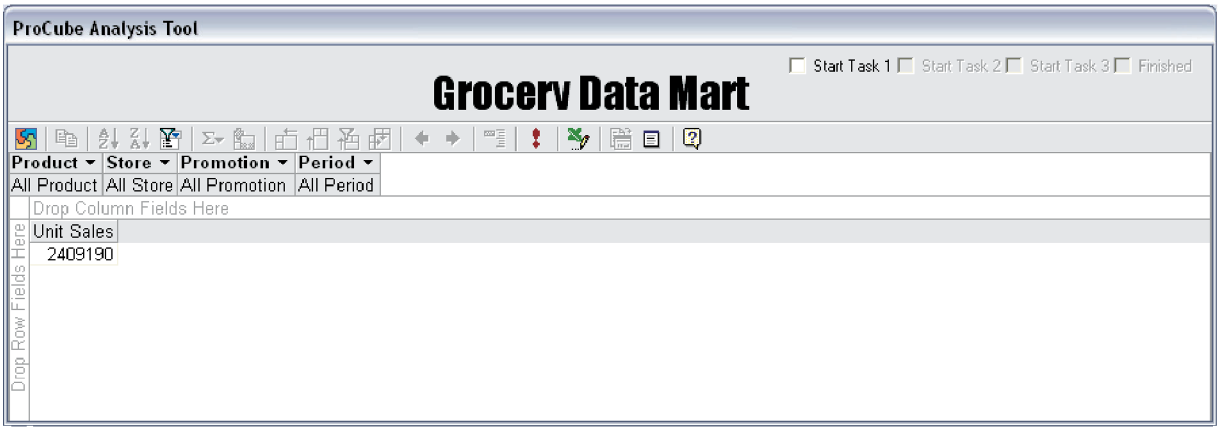

Figure 2: Pivot table based OLAP query interface used by the control group.

3. AN ALTERNATIVE TO THE PIVOT TABLE

Erik Thomsen, a leading figure in the OLAP community, has developed a simple tool to help designers of OLAP systems conceptualise the multi-dimensional structures they are creating. Originally, Thomsen called this these diagrams multi-dimensional domain structure diagrams (Thomsen, 1997). Later he renamed them multi-dimensional type structures (Thomsen, 2002). In this paper they will simply be called Thomsen diagrams. Figure 3a and 3b shows two representations of the same simple multi-dimensional data structure – for sales analysis in a grocery store (Kimball, 1996). In Figure 3a, a traditional star-schema (Kimball et al., 1999) is shown, in Figure 3b, a Thomsen diagram (Thomsen, 2002) based representation is shown.

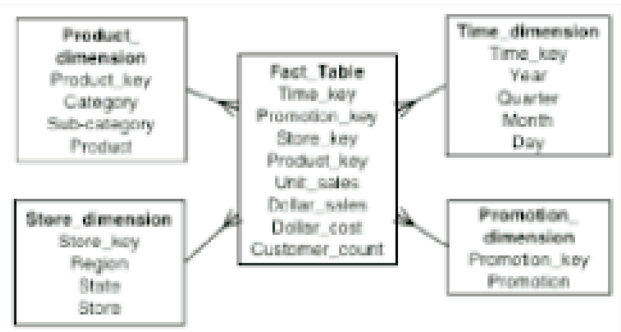

Figure 3a: Star-schema representation of the multi-dimensional grocery database.

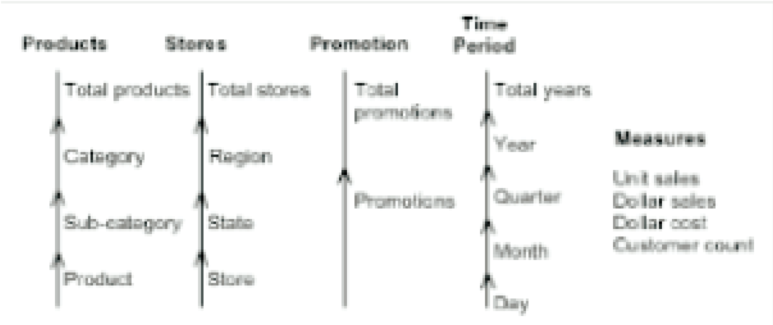

Figure 3b: Thomsen diagram representation of the multi-dimensional grocery database.

In a Thomsen diagram, the dimensions are shown as line segments (expect the measures dimension). The measures in the measure dimension are simply listed to the right of the line segments. Hierarchies within each dimension are explicitly displayed and the level labeled. No dimensions are treated specially, as row or column or page dimensions, except for the dimension that contains the measures – that is consistent with the physical reality of the OLAP data structure.

While Thomsen intended these diagrams be used by designers of OLAP structures, when thinking about the design of OLAP data structures, they are simple enough for end-users to understand. They can provide the basis of an interactive interface that can be used to access OLAP data. Figure 4 shows a screen capture of a prototype OLAP tool that uses an interface that is used is based upon a Thomsen diagram. The screen shown in Figure 4 is from the system used by the experimental group in the experiment described in the following sections of this paper. The system (available in Windows and Macintosh versions) can be downloaded from the web (URL to be advised). Users select the data that they require by interacting with the Thomsen diagram representation of the OLAP data structure. As fields are selected, the data displayed in the grid at the bottom of the screen is refreshed. A plain language description of the selected parameters and data is also generated automatically and displayed above the result data grid.

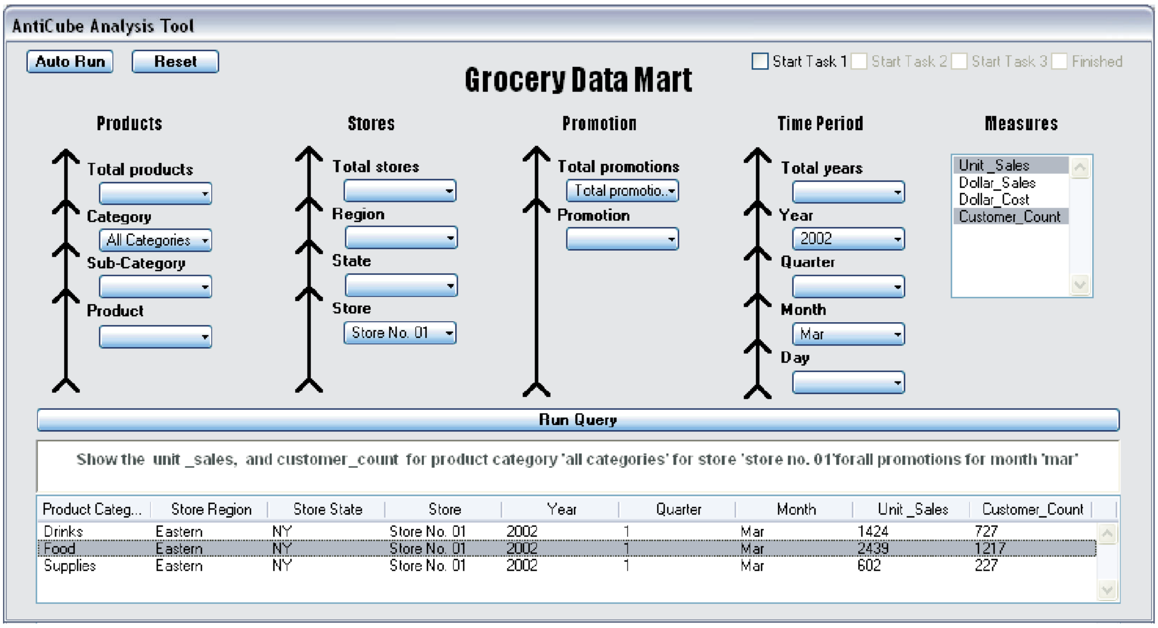

Figure 4: Thomsen diagram based prototype system interface.

4. EXPERIMENTAL METHOD

The experiment involved the interrogation of an OLAP database in a laboratory setting. One group of subjects was asked to interrogate an OLAP-based data set to answer a set of three questions using a software tool based on a pivot table. This group was the control group. Another group of subjects was asked to interrogate the same data set and answer the same set of questions. However this group, the experimental group, used a software tool with an interface based on the Thomsen diagram. The answers to the questions proved by the two groups were compared and analysed.

4.1 Experimental procedure

The subjects in the experiment were graduate students studying a unit on OLAP and Business Intelligence within a coursework Masters degree at an Australian University. These subjects should have had similar knowledge of the OLAP-based systems. Laboratory tutorials using a pivot table interface to access OLAP data and a study of Thomsen's diagramming method are included in the unit in which they were enrolled. At the end of the experimental procedure the subjects were also asked to complete a questionnaire designed to gauge their reaction to the experimental procedure. This included some data about the background of the subjects.

The experimental procedure was administered during a scheduled laboratory-based tutorial. Briefly the procedure was as follows: Each tutorial group was randomly allocated to be part of either the control or experimental group. The experimenter would meet the students at the beginning of the tutorial and explain the experiment, inviting the students to participate. Not all students in each tutorial group agreed to participate. Those who did agree to participate in the experiment were given a booklet and asked to read it. While they read the booklet the software tool they were to use was installed and set-up up on their computer. The booklet contained some information about the experiment, the database they were being asked to interrogate, and instructions on the use of the software tool including an illustration of the steps required to perform a task similar to the three experimental tasks. Once they finished reading the pre-amble in the booklet the subjects then read and performed each of the three experimental tasks, noting the answers to the tasks in the booklet as they went. When they finished the experimental tasks the software tool they were using then displayed a screen showing the time taken for each task. The subjects noted these times in their booklet. They then went on to answer the post-procedure questionnaire included at the end of the booklet. As each subject finished they were thanked for their participation by the experimenter who collected each booklet. Once finished, the subjects resumed their normal subject-based tutorial work in the laboratory.

4.2 Experimental tasks

Each subject was asked to perform three data analysis tasks. The tasks the subjects were asked to perform were intended to be typical of the types of data search and analysis that users of an OLAP system might be expected to perform. The tasks are listed in Table 1. Both the experimental and control groups were asked to perform the same tasks. The tasks involved interrogation of an OLAP database containing sales data for a grocery store. The database used was a simplified version of the sample dimensional data model provided by Ralph Kimball with his demonstration software tool Star-Tracker and discussed in his best selling data warehouse design text (Kimball, 1996). The date related data in this database was updated from the sample Kimball provides and the data was also altered so that there were unambiguous answers to the questions posed by the experimental tasks.

Task #	Description
Task 1	Marcus is the recently appointed marketing manager for a grocery chain. He has been informed that the company runs major promotions during the Christmas period, and his first duty is to implement an appropriate Christmas promotion. He would like to identify which promotion preformed best over the Christmas period in recent years, and implement it again this year. Which promotion should he implement and why?
Task 2	Jeannie is the Mid West sales manager for a major grocery chain. Recently there has been a dramatic drop in the overall sales for this region. She would like to identify which store is responsible for this drop in sales, and then take action accordingly. Which store is responsible for the drop in sales for the year of 2002? In which quarter did this sales drop occur?
Task 3	Lino is the inventory manager for 'Store 14' of a major grocery chain. Last year he under-stocked his store for the product category 'food', and was in turn reprimanded for it. He would like to avoid a repeat situation, by basing his order for the next year on some solid evidence. Therefore he would like to review the unit sales for the past years, prior to making his decision. Approximately how much should he order?

Table 1: Experimental tasks.

4.3 Measurement

Each subject wrote the answers to each of the analysis tasks in a section provided for them in the booklet that they were provided. When these were analysed, correct answers were scored 1 and incorrect answers scored 0. As a result, each subject received a score between 0 and 3 for their performance in completing the analysis tasks. As mentioned in the previous section, the database was configured so that there were unambiguous correct answers to each of the questions posed.

The software tool each subject used had a timer built in to record the length of time taken to complete each of the three tasks. When the subjects started, they clicked a button on screen to start the timer (note the top right of the screens shown in Figure 2 and 4). As the subjects completed each of the three tasks they clicked a marker to indicate that they had moved from one task to the next. When they finished the three tasks a window appeared showing the times taken for each of the tasks. The subject then recorded these times in their booklet before they moved to answer the post-procedure questions included in the booklet. The subjects could not close the window displaying the task completion times. Only the experimenter could close the window and shutdown the program using a secret key combination. This helped to ensure that the window wasn't accidentally closed and the times taken not recorded.

The workload of the subjects was measured using a subjective instrument based on the NASA task load index (NASA-TLX) (Hart & Staveland, 1988). The NASA-TLX is a well-known and widely used instrument for measuring subjective workload on subjects in experiments. It uses six dimensions to measure work activity: mental demand, physical demand, time pressure, performance, effort and frustration. A 20 step bi-polar scale is used to obtain ratings for these dimensions. The end points of each scale is *low* and *high* for each item except for performance which has end points *good* and *poor*. Subjects tick a point along the scale, which is then coded as a number from 1 to 20. Usually the ratings for each dimension are weighted and combined to give an overall score for subjective workload. These weights are developed using a long set of paired comparison tasks. For this experiment it decided not to combine the dimensions in this way. This enabled the total time taken by the subjects to complete the tasks to be kept to a minimum – eliminating the need for the paired comparisons to be completed by the subjects. In this experiment the item physical demand was also dropped from the instrument leaving five items. It was felt that this item was irrelevant to the software tools being investigated.

4.4 Hypothesis

Hypothesis 1 examines the accuracy of the answers the subjects in the two groups gave to the questions they were asked to answer in the three experimental tasks. Hypothesis 2 examines the time taken to complete all of the tasks by the two groups of subjects. Each of the hypotheses 3 to 7 examines one of the five dimensions of work load as rated by the subjects in each of the groups.

Hypothesis 1: Users of the Thomsen diagram based interface will get more correct answers to the questions posed by the three experimental tasks than will users of the pivot table based interface.

Hypothesis 2: Users of the Thomsen diagram based interface will complete the three experimental tasks in less time than the users of the pivot table based interface.

Hypothesis 3: Users of the Thomsen diagram based interface will experience a lower mental demand when completing the three experimental tasks in less time than the users of the pivot table based interface.

Hypothesis 4: Users of the Thomsen diagram based interface will require less effort when completing the three experimental tasks the users of the pivot table based interface.

Hypothesis 5: Users of the Thomsen diagram based interface will perceive less time pressure to complete the three experimental tasks than the users of the pivot table based interface.

Hypothesis 6: Users of the Thomsen diagram based interface will complete the three experimental tasks in less time than the users of the pivot table based interface.

Hypothesis 7: Users of the Thomsen diagram based interface will report a lower frustration level than the users of the pivot table based interface.

5. RESULTS

5.1 Demographics

A total of 50 students participated in the experiment with an even allocation of 25 subjects to the control and experimental groups. 33 of the subjects were male and 17 were female. Just over half of the subjects belonged to 21-24 age group, less than half belonged to the 25 and over age group, with only one participant aged less than 21. All the subjects for this experiment were graduate students undertaking an elective unit in OLAP and Business Intelligence.

The subjects were asked a series of questions regarding their background. They were asked about their knowledge of and experience with OLAP tools. The majority of subjects had gained their knowledge of OLAP from their study in the unit they were enrolled in. A small number, 8% (2 out of 25) of the experimental group and 12% (3 out of 25) of the control group, had some OLAP knowledge gained from their professional work. Similarly, only a small number 8% (2 out of 25) of the experimental group and 16% (4 out of 25) of the control group had used OLAP tools outside of their studies.

The subjects were asked further background question concerning their computer skills. The subjects were asked to rate their computer skill levels as either being 'high', 'medium' or 'low'. For both groups, the majority possessed either a 'high' or 'medium' level, with 40% (10 of 25) and 48% (12 of 25) respectively belonging to the experimental group and 36% (9 of 25) and 52% (13 of 25) respectively part of the control group. A number of the subjects did not provided a response, 12% (3 of 25) from the experimental and 8% (2 of 25) from the control. This data confirms the expectation that graduate information technology students would have significant skills interacting with computers.

5.2 Summary of results

Neither group answered the questions associated with the three experimental tasks very well. In the experimental group over half got at least 1 answer wrong with only 48% getting every answer correct. The control group did worse with a strong majority (72%) getting at least one answer wrong with only 28% of subjects getting all the answers correct. Only 1 subject in the experimental group got all the answers wrong. In the control group nearly a third of the subjects got all the answers wrong. The counts of the number of correct answers for the experimental group and the control group are shown in Table 2.

There was very little difference between the experimental group and the control group for the mean total time to complete the experimental tasks. The mean time taken by the experimental group to complete all three tasks was 15 minutes 52 seconds (std dev. 5 min 41 sec, n 23 – two subjects in this group did not record their task completion times). The mean time taken by the control group was 16 minutes 27 seconds (std dev. 7 min 37 sec, n 25).

Participants in both groups experienced similar levels of mental demand, perceived success and frustration. All of these variables were rated lower than the mid-point of the 1-20 scale used (10.5). Levels of time pressure and computational effort required differed slightly. The mean and standard deviation for each of the subjective workload measures is shown in Table 3. A simple t-test (Snedecor & Cochran, 1989) was performed for each of the variables for each group to compare the mean value for each group to this mid-point. This test indicates whether the difference between the mean score and the mid-point was due to chance.

Number of correct answers	Experimental group	Control group
3	12 (48%)	7 (28%)
2	7 (28%)	6 (24%)
1	5 (20%)	4 (16%)
0	1 (4%)	8 (32%)
Total	25 (100%)	25 (100%)

Table 2: Number of correct answers for the experimental and control groups.

Mental demand	**n**	**Mean**	**Std Dev**	**T**
Experimental	24	7.83	3.54	-3.69*
Control	24	7.79	3.64	-3.64*
Effort required	**n**	**Mean**	**Std Dev**	**T**
Experimental	24	7.92	4.31	-2.93*
Control	24	9.68	4.18	-0.98
Perceived success	**n**	**Mean**	**Std Dev**	**T**
Experimental	24	12.38	3.95	2.32*
Control	25	12.16	4.80	1.72
Time pressure	**n**	**Mean**	**Std Dev**	**T**
Experimental	24	6.75	5.34	-3.43*
Control	25	8.68	4.35	-2.09*
Frustration level	**n**	**Mean**	**Std Dev**	**T**
Experimental	24	9.00	3.50	-0.77
Control	25	9.70	5.18	-2.09*

Note: "T" column shows the value of t for a one-sample t-test comparing results to the neutral mid-point of the scale. Values of t that show that the difference between the mean and the neutral mid-point of the scale (10.5) is not due to chance (significant level 95%) are marked by a "*".

Table 3: Results of the subjective workload measures the experimental and control groups.

For the questions regarding the experiment and the prototype, subjects stated their answer as either 'yes, 'no' or 'indifferent'. The majority of participants stated that they enjoyed the experiment, with 80% (20 of 25) from the experimental group and 88% (22 of 25) from the control group answering 'yes'. The majority - 80% (20 of 25) - of subjects within the experimental group enjoyed using the program. However, fewer - only 68% (17 of 25) - subjects in the control group stated 'yes' to this question, with the 20% (5 of 25) indifferent regarding the prototype.

5.3 Hypothesis testing

The test used to perform the testing of hypothesis 1 was the rank sum test (Mann & Whitney, 1947; Wilcoxon, 1945) . This test uses a z-score to determine if the difference between two non-parametric data sets can be attributed to chance. The critical value of z for a one-tailed hypothesis test at a 95% level of confidence is 1.96. The calculations associated with the rank sum test for hypothesis 1 are given in table 4. This table shows that, with a z-score of 2.12 (p = 0.034), the null form of hypothesis 1 can be rejected and that there is support for the alternative form. That is users of the Thomsen diagram based query tool got more correct answers to the questions posed by the three experimental tasks than did users of the pivot table based interface.

The other hypotheses (number 2 to 7) were tested using t-tests (Snedecor & Cochran, 1989). Levene's test for equality of variances was used to ensure the assumption of equal variances was valid for each hypothesis (Levene, 1960). For each hypothesis, except for hypothesis 6, this assumption was found to be valid and the standard form of the t-test that assumes equal variances was used. For hypothesis 6, the Welch-Satterthwaite form of the t-test that doesn't assume equal variances was used. The calculations and results of those tests are shown in Table 5. Using a significance level of 0.05 (95%) there is no support for the alternative form of hypotheses 2 through 7. For each of these hypotheses the null form must be accepted – that is there is no difference between the experimental and control groups for the time taken to complete the tasks, the perceived mental demand, the perceived effort, perceived success, time pressure and for frustration level.

Number correct	Frequency (Experimental Group)	Frequency (Control Group)	Ranks	Mean rank	Frequency by mean rank (Experimental Group)	Frequency by mean rank (Control Group)
3	12	7	32-50	41	492	287
2	7	6	19-31	25	175	150
1	5	4	10-18	14	70	56
0	1	8	1-9	5	5	40
	$N_{Control}$ = 25	$N_{Experimental}$ = 25			$T_{Experimental}$= 742	$T_{Control}$ = 533

Wilcoxon W	533
Mann-Whitney U	208
Z-score (adjusted for ties)	2.12

Table 4: Calculations for the rank sum test used to test hypothesis 1

	Experimental group			Control group					
Hypothesis	Mean	Std dev.	n	Mean	Std dev.	n	df	t	p
2: Time taken	15:52	05:41.19	23	16:27	07:36.65	25	46	0.305	0.381
3: Mental demand	7.83	3.54	24	8.0	3.8	24	46	-0.040	0.484
4: Effort	7.92	4.31	24	9.7	4.2	25	47	1.453	0.076
5: Perceived success	12.38	3.95	24	12.2	4.8	25	47	-0.171	0.432
6: Time pressure	6.75	5.34	24	8.7	4.3	25	47	1.384	0.086
7: Frustration level	9.00	3.50	24	9.7	5.2	25	47	0.556	0.290

Table 5: *T*-tests for hypotheses 2 to 7

5.4 Discussion

5.4.1 Findings

The most dramatic and important result of this project is the support found for hypothesis 1. That is that subjects using the Thomsen diagram-based interface performed better at the analysis tasks than the subjects that used the pivot table. The performance of both groups was quite poor, however, for the group that used the pivot table they were simply terrible. It is worth noting that the subjects in this experiment had some experience using a pivot table and while they were familiar with Thomsen diagrams they had not had any prior experience using a software tool based on a Thomsen diagram. It would have been reasonable to assume that this may have biased the results in favour of the group using pivot table. This reinforces the proposition that the pivot table – a widely used interface component in business intelligence systems – is difficult to use, and that it can and should be improved upon. The alternative OLAP interface proposed here, whilst more effective than the pivot table, is only one alternative. Others could be developed that may prove to be even more effective.

While there was no support for the other hypotheses, the results related to these hypotheses are interesting. The results do not show that the Thomsen diagram based tool was any better than the pivot table-based tool on these other measures of task performance (time taken, mental demands, effort, perceived success, time pressure and frustration level), however, the results do show that the alternative tool is at least equivalent to a pivot table for these performance criteria. For some of these criteria there is some effect, though not significant. For effort, perceived success and possibly also for frustration level a larger sample size, increasing the statistical power of the analysis may yield a significant result.

5.4.2 Study limitations

It must be noted that the study is subject to several limitations that may limit the generalisability of the findings. First, graduate students with little 'real' experience of OLAP analysis were used. However, end-users of OLAP applications would not normally be expected to have experience doing analysis with that OLAP technology. Another limitation relates to the used of 'canned' tasks. The tasks, whilst realistic, were not real. The subjects had no personal interest in solving the problems. They were not rewarded in any way for participation in the experiment nor was their performance rewarded or punished in any way. The use of subjects working with a data set related to their actual work on problems that they have a personal interest in might lead to better performance and fewer errors. The laboratory environment in which the experiment was conducted is also a little removed from the real environment in which people use OLAP tools. In their own office, a typical OLAP user has other people to talk to and ask for help, manuals to refer to and perhaps access to old analyses - parts of which they might be able to reuse.

Further, the measures used to gauge the performance of the subjects in this experiment record the outcome of their interaction with the software tools while performing the analysis tasks. They do not provide any information about the way these tools were used. For example the mistakes or misunderstanding the users made that lead them to get incorrect answers to the questions they were asked. Further research is required to better understand the actual cognition that is taking place when a user queries an OLAP data structure and how the interface of the query tool supports or hinders their analysis.

6. CONCLUDING COMMENTS

Over the last few years the pivot table has become the standard interface for access to OLAP-based data. This study has shown that the pivot table is not a very effective tool for *ad hoc* access to OLAP data. In the experimental study described here users of a pivot table performed badly, with the majority being unable to correctly perform typical simple OLAP analysis tasks. An alternative interface, based on the diagramming technique developed by Erik Thomsen for the design of OLAP data structures, has been developed. This interface has – in this study – been shown to be a more effective tool for typical OLAP analyses than the traditional and widely used pivot table.

OLAP technology is often referred to as a technology that helps users visualise their business. In that case, the interface provided to users needs to aid that visualisation and assist in their exploration of data. The interface to OLAP data must be easily understood, easy to learn and relatively intuitive. Currently, there is little to distinguish the pivot table style interface offered by the vendors in the OLAP client market. The vendor who is able to develop a more effective interface will obtain an important competitive and market advantage. Unless this happens, the use of OLAP in business intelligence systems will be restricted to 'power-users'. As a result OLAP-based systems may not contribute to improvements in the design of business intelligence systems and to improving corporate decision-making.

REFERENCES

Codd, E. F., Codd, S. B., & Salley, C. T. (1993). *Providing OLAP to User Analysts: An IT Mandate* (Web page): E.F. Codd and Associates.

Crandall, R. (1996). Foreword. In E. Thomsen (Ed.), *OLAP Solutions: Building Multidimensional Information Systems* (pp. xv-xxv): John Wiley.

EMC, Unisys, & Microsoft. (2002). *T3 Business Intelligence Project: An Enterprise Server and Storage Architecture for Scalability, Availability, and Manageability* (White paper): EMC.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*: North-Holland: Elsevier.

Kimball, R. (1996). *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*: John Wiley.

Kimball, R., Reeves, L., Ross , M., & Thornwaite, W. (1999). *The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing and Deploying Data Warehouses*: John Wiley.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278-292): Stanford University Press.

Mann, H. B., & Whitney, D. R. (1947). On a test whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist., 18*, 50-60.

Martin, C. J., & Clarke, B. A. (1990). Executive Information Systems: Recent Developments and Research Imperatives. *British Journal of Management, 1*, 27-35.

McKendrick, J. (1998). Microsoft adds a new dimension to the OLAP market. *ENT Magazine*.

Microsoft. (2003). Office XP BI Portal (Version 2002): Microsoft.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods* (Eighth ed.): Iowa State University Press.

Thomsen, E. (1997). *OLAP Solutions: Building Multidimensional Information Systems* (1st ed.): John Wiley.

Thomsen, E. (2002). *OLAP Solutions: Building Multidimensional Information Systems* (2nd ed.): John Wiley.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics, 1*, 80-83.

XMLA. (2003). Interoperability Data Set - "The Beverage Company" (Version 4) [download from http://www.xmla.org/download.asp?id=56]: XML for Analysis Council.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the useful feedback from the anonymous reviewers, whose contributions have improved the ongoing work involved in this project.

COPYRIGHT