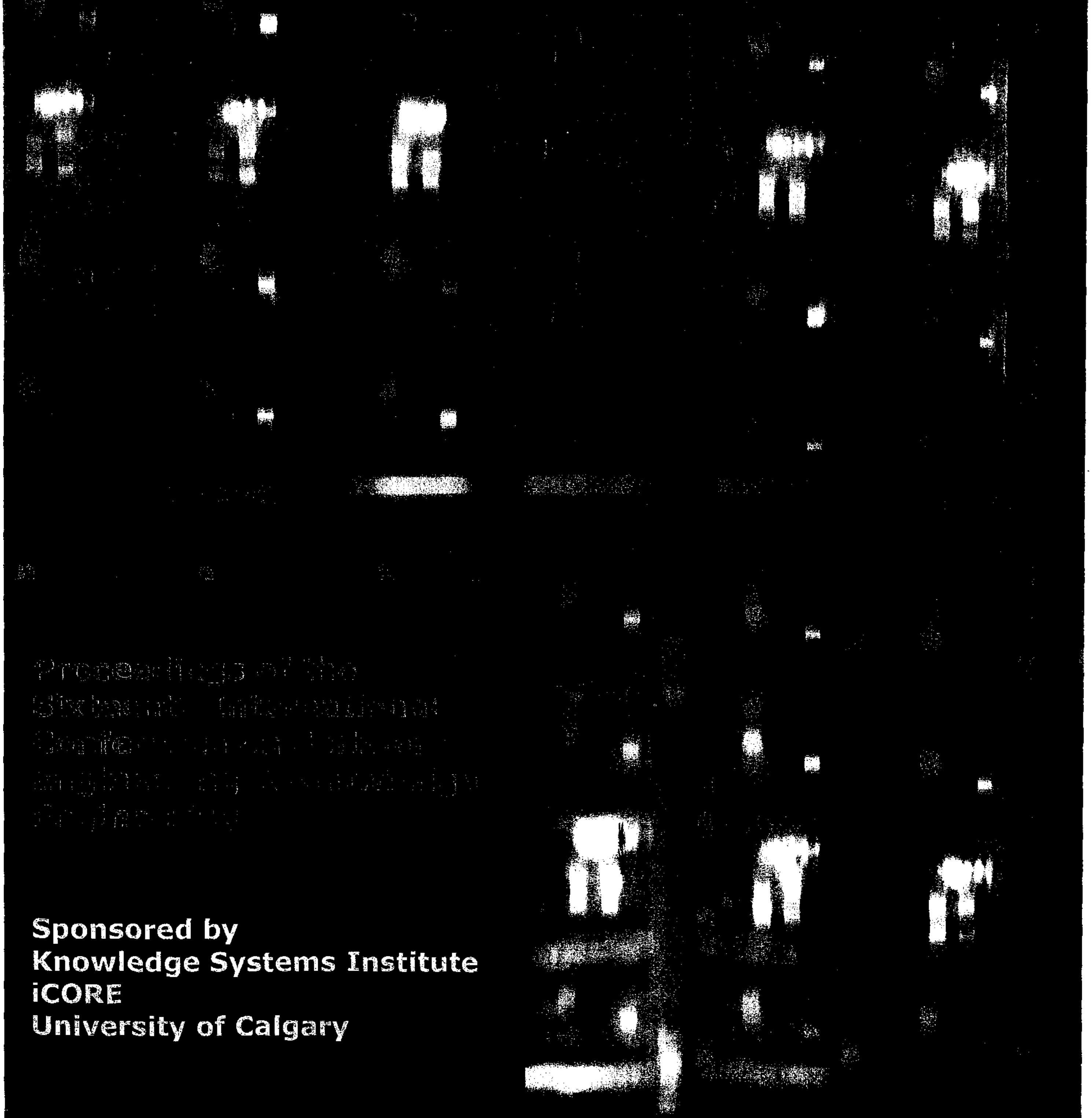
2004 SEKE

Alberta, Canada
June 20 to June 24, 2004



Proceedings of the
Sixteenth International
Conference on Software
Engineering & Knowledge
Engineering

Sponsored by
Knowledge Systems Institute
iCORE
University of Calgary

PROCEEDINGS

SEKE 2004

The 16th International Conference on Software Engineering & Knowledge Engineering

Sponsored by

Knowledge Systems Institute Graduate School, USA

Co-Sponsored by

Informatics Circle of Research Excellence, Canada
University of Calgary, Canada

Technical Program
June 20-24, 2004
Banff Alberta, Canada

Organized by

Knowledge Systems Institute Graduate School

[redacted]

ISBN 1-891706-14-4 (paper)

Additional Copies can be ordered from:
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076, USA
Tel:+1-847-679-3135
Fax:+1-847-679-3166
Email:office@ksi.edu
http://www.ksi.edu

Printed in the United States of America

The 16th International Conference on Software Engineering & Knowledge Engineering (SEKE'2004)

June 20-24, 2004
Banff, Alberta, Canada

Organizers & Committee

General Chair

Shi-Kuo Chang, Univ. of Pittsburgh, USA

Program Co-Chairs

Frank Mauer, Associate Head, Department of Computer Science, Univ. of Calgary, Canada
Guenther Ruhe, Industrial Research Chair in Software Engineering, Univ. of Calgary, Canada

Publicity Chair

Jens H. Jahnke, Univ. of Victoria, Canada

Program Committee

Silvia Teresita Acuna, Univ. Autónoma de Madrid, Spain
Anneliese K. Amschler Andrews, Washington State Univ., US
Juan Carlos Augusto, Univ. of Ulster,UK
Aybuke Aurum, Univ. of New South Wales, Australia
Frank Bomarius, Fraunhofer IESE, Germany
Paolo Ciancarini, Univ. of Bologna, Italy
William Chu, Tunghai Univ., Taiwan
John Debenham, Univ. of Technology, Australia
Andrea De Lucia, Univ. of Salerno, Italy
Yi Deng, Florida International Univ., US
Schahram Dustdar, Vienna Univ. of Technology, Austria
Haka Erdogmus, National Research Council Canada, Canada
Filomena Ferrucci, Univ. of Salerno, Italy
Alfonso Fuggetta, Politechnico di Milano Technical Univ.y, Italy
Carlo Ghezzi, Politechnico di Milano Technical Univ., Italy

Athula Ginige, Univ. of Western Sydney, Australia
Christiane Gresse von Wangenheim, Universidade do Vale do Itajaí, Brazil
Volker Gruhn, Univ. of Leipzig, Gemany
John Grundy, Univ. of Auckland, New Zealand
Mao Lin Huang, Univ. of Technology, Australia
Hajimu Iida, Nara Institute of Science and Technology, Japan
Letizia Jaccheri, Norwegian Univ. of Science and Technology, Norway
Natalia Juristo, Technical Univ. of Madrid, Spain
Huimin Lin, Chinese Academy of Sciences, China
Mikael Lindvall, Fraunhofer Center Maryland, US
Jiming Liu, Hong Kong Baptists Univ., China
Luqi, Naval Postgraduate School, US
Sandro Morasca, Univ. degli Studi dell'Insubria, Italy
Jurgen Munch, Fraunhofer IESE, Germany
Lakshmi Narasimhan, Univ. of Newcastle, Australia
Paolo Nesi, Univ. of Florence, Italy
Mehmet Orgun, Macquariez Univ., Australia
Michael Richter, Universitat Kaiserslautern, Germany
Ioana Rus, Fraunhofer IESE, US
Walt Scacchi, Univ. of California Irvine, US
Phillip Sheu, Univ. of California Irvine, US
Eleni Stroulia, Univ. of Alberta, Canada
Scott Tilley, Florida Institute of Technology, US
Genny Tortora, Univ. of Salerno, Italy
Jeffrey Tsai, Univ. of California Irvine, US
Sira Vegas, Universidad Politécnica de Madrid, Spain
Giuseppe Visaggio, Univ. of Bari, Italy
Giuliana Vitiello, Univ. of Salerno, Italy
Yingxu Wang, Univ. of Calgary, Canada
Stefan Wermter, Univ. of Sunderland, UK
Xindong Wu, Univ. of Vermont, US
Yiyu Yao, Univ. of Regina, Canada
Kang Zhang, Univ. of Texas, US

Proceedings Cover Designer

Gabriel Smith, Knowledge Systems Institute Graduate School, USA

Conference Secretariat

Judy Pan, Chair, Knowledge Systems Institute Graduate School, USA
Tony Gong, Knowledge Systems Institute Graduate School, USA
C. C. Huang, Knowledge Systems Institute Graduate School, USA
Beverly Huggins, Knowledge Systems Institute Graduate School, USA
Rex Lee, Knowledge Systems Institute Graduate School, USA
Daniel Li, Knowledge Systems Institute Graduate School, USA

International Workshop on Knowledge Oriented Maintenance (KOM'2004)

June 20, 2004
Banff, Alberta, Canada

Organizers

Nicolas Anquetil, Department of Knowledge Management and Information Technology (MGCTI), Catholic Univ. of Brasília, Brazil
Timothy C. Lethbridge, School of Information Technology and Engineering, Univ. of Ottawa, Canada

Program Committee

Nicolas Anquetil, Brazil
Giuliano Antoniol, Italy
Françoise Balmas, France
Dirk Deridder, Belgium
Nicolas Gold, United Kingdom
Idris H. His, U.S.A.
Timothy C. Lethbridge, Canada
Andrea de Lucia, Italy
Ettore Merlo, Canada
Rosângela Ap. Dellosso Penteado, Brazil
Eleni Stroulia, Canada

[redacted]

International Workshop on Ontology In Action (OIA'2004)

June 21, 2004
Banff, Alberta, Canada

Workshop Co-Chairs

Athula Ginige, School of Computing and Information Technology, Univ. of Western Sydney, Australia
Káthia Marçal de Oliveira, Catholic Univ. of Brasília, Brasília, DF, Brazil

Program Committee

Bernhard Holtkamp, Fraunhofer Institut Software-und Systemtechnik, Germany
Ricardo de Almeida Falbo, Federal Univ. of Espírito Santo, Brazil
John Domingue, Knowledge Media Institute, The Open Univ., UK
Uma Srinivasan, CSIRO ICT Centre, Australia
Steffen Staab, Univ. of Karlsruhe, Germany
Kerry Taylor, CSIRO ICT Centre, Australia
Giancarlo Guizzardi, Univ. of Twente, The Netherlands
Luigi Ceccaroni, Univ. Politecnica de Catalunya, Spain
Germana Meneses da Nóbrega, Catholic Univ. of Brasília, Brazil
Mike Uschold, The Boeing Company, Seattle, USA

3rd International Workshop on Software Engineering Decision Support (SEDECS'2004)

June 22, 2004
Banff, Alberta, Canada

Program Chair

Guenther Ruhe, Industrial Research Chair in Software Engineering, Univ. of Calgary, Canada

Program Committee

Stephan Biffl, Austria
Khaled El-Emam, Canada
Ross Jeffery, Australia
Sandro Morasca, Italy
Dietmar Pfahl, Germany
David Raffo, USA
Iona Rus, USA
Claes Wohlin, Sweden

Foreword

The Sixteenth International Conference on Software Engineering and Knowledge Engineering (SEKE 2004) will be held at the Banff Centre, Banff, Alberta, Canada from June 20 to June 24, 2004. The conference brings together experts in software engineering and knowledge engineering to discuss relevant results in both disciplines. Special emphasis is put on synergies between both domains. The conference received nearly 150 technical papers. After a detailed review process, 38% of the submissions were accepted as long papers and an additional 17% as short presentations. Long papers were accepted based on their research quality while short papers and workshop submissions usually report on research in progress and new ideas.

The conference presentations cover a wide spectrum of software engineering and knowledge engineering topics including software processes and process improvement, experience management, quality assurance & testing, requirements engineering, decision support and fuzzy SE knowledge, web engineering, ontologies and agent technology, design and patterns, and formal specification. Authors provide new insights and perspectives on future research directions. The papers included in the conference proceedings speak for themselves.

Several workshops are running in addition to the main conference. The Canadian Agile Network was invited to take part in its Second Canadian Agile Network Workshop. The goal of the workshop is to disseminate ideas, lessons learned and best practices of adopting agile methods and moving them to the mainstream of software development. This year the main focus is on agile culture, following organizational change and agile project management.

The Workshop on Knowledge-oriented Maintenance investigates the role of "knowledge" on software maintenance process. Participants share their experience on the extraction and application of knowledge in software maintenance processes.

The Workshop on Learning Software Organizations (LSO 2004) brings together researchers from industry and academia to discuss how continuous learning processes can be implemented and supported in software development teams. Its focus is on practical applications and experience reports.

The Workshop "Ontology in Action" elaborates on how shared ontologies can be formalized and used for sharing information among heterogeneous software applications. It focuses on how semantic interoperability can be reached by modeling entities and their relationships as domain ontologies.

The Workshop on Software Engineering Decision Support is devoted to discuss methodology, tools and experience on providing support for decision-making in software development and evolution.

We are grateful to all the members of the Program Committee: Silvia Teresita Acuna, Anneliese K. Amschler Andrews, Juan Carlos Augusto, Aybuke Aurum, Frank Bomarius, Paolo Ciancarini, William Chu, John Debenham, Andrea De Lucia, Yi Deng, Schahram Dustdar, Haka Erdogmus, Filomena Ferrucci, Alfonso Fuggetta, Carlo Ghezzi, Athula Ginige, Christiane Gresse von Wangenheim, Volker Gruhn, John Grundy, Mao Lin Huang, Hajimu Iida, Letizia Jaccheri, Natalia Juristo, Huimin Lin , Mikael Lindvall, Jiming Liu, Luqi, Sandro Morasca, Juergen Muench, Lakshmi Narasimhan, Paolo Nesi, Mehmet Orgun, Michael Richter, Ioana Rus, Walt Scacchi, Phillip Sheu, Eleni Stroulia, Scott Tilley, Genny Tortora, Jeffrey Tsai, Sira Vegas, Giuseppe Visaggio, Giuliana Vitiello, Yingxu Wang, Stefan Wermter, Xindong Wu, Yiyu Yao, Kang Zhang. The program committee did an enormous job to review a large number of submitted papers. Their effort ensured the final quality of the conference and all the workshops.

In addition to our program committee members, we would like to thank the following reviewers for providing feedback on submitted papers: Piefrancesco Bellini, Sami Beydeda, Alessandro Bianchi, Kai-Yuan Cai, Zhining Cao, Rosa M. Carro, Alejandra Cechich, María Dolores Vargas Cerdán, Yurong Chen, Oscar Corcho, Patricia Costa, Feras T. Dabous, Angélica de Antonio, Vincenzo Deufemia, Oscar Dieste, Paolo Donzelli, Toncan Duong, Pascal Fenkam, Xavier Ferre, Andres Flores, Rita Francese, Shu Gao, Marisol Giardina, Haitao Gong, Carmine Gravino, Thomas Gschwind, Mariele Hagen, Aaron Hector, Bayu Hendradjaya, Pilar Herrero, Lorin Hochstein, Siv Hilde Houmb, Hiroshi Igaki, José Antonio Macías Iglesias, Zhi Jin, Kanta Jiwnani, André Köhler, Jun Kong, Serguei Krivov, Cat Kutay, Guojun Li, Jingzhou Li, Sheldon X. Liang, Hong-Xin Lin, Pdero Linares, Fabiola Lopez y Lopez, Sergio Di Martino, Nelson Medinlla, Gonzalo Méndez, Abdallah Mohamed, Ana M. Moreno, Ming Muo, Abhaya Nayak, Josef Nedstam, An Ngo-The, Andrew O'Fallon, Alvaro Ortigosa, Thomas Østerlie, Luca Paolino, Orest Pilskalns, Martin Pinzger, Giuseppe Polese, Yu Qian, Fethi Rabhi, Jaime Ramírez, Michele Risi, Omolade Saliu, Marisa Sanchez, Maria-Isabel Sanchez-Segura, Maribel Sanchez-Segura, Giuseppe Scanniello, Klaus Schmid, Indra Seher, Michele A. Shaw, John Shepherd, Xiaochun Shi, Alejandro Sierra, Almudena Sierra-Alonso, Janice Singer, Guanglei Song, Lorna Stewart, Weixiang Sun, Magne Syrstad, Cora B. Excelente Toledo, Maximiliano Paredes Velasco, Qing Wang, Richard Webber, Ying Yang, JingTao Yao, Huilin Ye, InSeon Yoo, Huiqun Yu, Hairong Yu, Guangcun Zhang, Xu Zhang, Haiyan Zhao, Liming Zhu, Xingquan Zhu

Special thanks to all the sponsors of the conference: The Informatics Circle of Research Excellence (iCORE), the University of Calgary and the Knowledge Systems Institute Graduate School.

Welcome to SEKE'2004!

Frank Maurer & Guenther Ruhe
SEKE 2004 Program Committee Co-Chairs

Table of Contents

Keynote Papers

SEKE Long Papers

SEKE Short Papers

KOM Workshop Papers

OIA Workshop Papers

SEDECS Workshop Papers

Datawarehouses design: effectivity of the star schema

Coral Calero, Manuel Serrano, Mario Piattini
ALARCOS Research Group
Department of Computer Science-University of Castilla-La Mancha
Paseo de la Universidad, 4 13071 Ciudad Real (Spain)
{Coral.Calero, Manuel.Serrano, Mario.Piattini}@uclm.es

Abstract. Datawarehouses have become the most important tread in business, and it is essential that the design be made to assure efficiency and simplicity of use. Although it is generally accepted that the star design is the best way to implement datawarehouses in relational database management systems there are no studies that confirm this assumption. With the aim of determining if the star design really gives more comprehensible datawarehouses, we are carrying out a series of experiments. In this article we present the studies done up to now showing that there is no difference in difficulty when using the "traditional" relational method (using E/R modeling and then transforming it into the relational model) than when using the star design.

1. Introduction

Datawarehouses arose due to the need of organizations to have mechanisms that helped in decision making. Datawarehouses have become the most important trend in organizations since they provide information for improving strategic decisions. Datawarehouses and related business intelligence technnologies have increased enormously in recent years, and are expected to reach 150 billion dollars in 2005 (Jarke et al., 2000; Chenoweth et al., 2003).

If the datawarehouse has been properly constructed, it provides the organization with a foundation that is extremely flexible and reusable (Inmon, 2002). So, it is essential to assure that the datawarehouse is designed properly. One way of designing a datawarehouse is to use dimensional modelling, a logical design technique alternative to the classical database design based on entity-relationship modelling (together with the transformation to the relational model).

The dimensional modelling technique seeks to present the data in an intuitive standard framework that allows high-performance access. The dimensional model is composed of a table called the fact table (the primary table that is meant to contain measurements of the business) and a set of smaller tables called dimensional tables.

Some authors, such as Kimball et al (1998) argue that dimensional modelling is the only viable technique for delivering data to end users in a datawarehouse and has a number of advantages over the entity-relationship based methodology. In general, it is argued that decision-oriented dimensional datawarehouses are fundamentally different from transaction-oriented relational databases and a different set of tools is required for their effective development (Chenoweth et al., 2003).

There is, however, no proof of this assumption. In an attempt to clarify this assumption, we decided to do a series of experiments with the intention demonstrating that the use of the star model makes datawarehouses simpler to use than when the traditional design is used (relational and E/R).

The work done until now can be divided into two studies (one of them was replicated twice). In this article we present all the results that we have obtained from all these experimental studies.

In the following section all the experimental processes of each of the studies together with its conclusions are described. Section three presents the conclusions and describes future studies.

2. Experimental work

The objective of our work is to determine if it is better to design the datawarehouse using a traditional design methodology (that begins by model E/R and follows with the transformation to the relational model) or by using the star design. Until now we have carried out two experiments and we have replicated one of them twice.

In both experiments the working hypotheses, the dependent and independent variables with which the experiments work, the experimental material, the execution process and many of the threats to the validity of the experiments are the same. Before explaining in depth each experiment and the results obtained we will explain in detail the common elements.

2.1. Hypotheses.

Our hypotheses are:
Null hypothesis (H0): There is no difference in the understandability of the two kinds of schema (traditional and star).
Alternative hypothesis (H1): There is a difference in the understandability of the two kinds of schema (traditional and star).

2.2. Dependent and independent variables.

In every case we have measured the dependent variable (understandability) by means of the time used by each subject in making the indicated tasks, the independent variable being the model used for the logical representation of the datawarehouse (traditional or star).

2.3. Experimental material and execution

The experimental material and the form in which the experiments were carried out were similar in both experiments. The schematas were provided to the subjects (six in the first study and twelve in the second) together with the questions. The subjects must indicate how and to what tables of each schema they must gain access to recover certain information from the datawarehouse. They also had to write down the time used in responding to each of the questions.
The experiment was done in one session. Before carrying out the experiment we gave an intensive explanation of what kind of problems they had to solve, how to answer the questions and what sort of material was provided for the accomplishment of the experiment. In no case did the subjects have knowledge of the aspects that we were trying to study or of the hypotheses we were working with.

2.4. Threats to validity

In order to be able to avoid diverse threats to the validity of the experiment, we tried to control some aspects:

- In each experiment subjects had similar experience and knowledge. Although to work with students might not appear rigorous, there exist studies that affirm that the differences between students and professionals are small and studies with students are viable under certain conditions (Hörst et al., 2000).
- The domains of the diagrams were simple and sufficiently known to avoid problems of understanding.
- In order to avoid learning effects the schematas were given to each subject in a different order.
- As it was the first time that the subjects performed an experiment of this type, persistence effects do not exist.
- Subjects were motivated because the exercises were included into the knowledge they had to acquire in their training.
- Plagiarism was controlled and conversations were not allowed among subjects during the experiment.
- All the doubts were solved by the person who led the experiment.

In the following two sections we will present in depth individual aspects of each of the experimental studies.

3. First experiment

As we have already mentioned this first experiment was replicatd twice so, finally we had three examples of the same experiment (an original experiment and two replicas).

3.1. Experimental design

The experiment consisted of six schematas, three traditional schemes and three semantically equivalent schemas designed using star diagrams. Subjects had to formulate SQL queries and to write down the time (in seconds) that it took make them.

3.2. Subjects

In the original experiment the subjects were 11 PhD students at the School of Computer Science of the University of Castilla-La Mancha (UCLM) in Spain. One of the replicas was made by 12 undergraduate students of the UCLM who were enrolled in the final year (when the experiment was done they were following a course on databases lasting two semesters) whereas the other replica was carried out with 24 PhD students of the Pinar del Rio University (Cuba).
In all the cases the subjects had knowledge of design and use of databases and datawarehouses. In addition the PhD students had a deeper knowledge of datawarehouses because they had received this information as part of their PhD studies.

3.3. Limitations

We are conscious of some limitations associated with these experiments such as the small number of subjects and objects and the difficulty associated with the use of SQL for the specification of the exercises given to the subjects.

3.4. Results

For our experiment we fixed a value $\alpha = 0.1$ to increase the power of the statistical tests (that is to say, the probability of rejecting our hypotheses when these are false). Due to the experimental design and to the gathered data the most suitable test is a repeated measure univariate ANOVA test (SPSS 11, 2001).

In tables 1, 2 and 3 the results obtained after applying the statistical test to the data gathered from the experiment are shown. Analyzing the significance value we can observe that all the values are greater than α and, therefore, we cannot reject the null hypothesis. So, there is no difference in the time used to answer the questions depending on the type of design used (Schema_type variable).

Table 1. ANOVA results for PhD students from UCLM (Spain)

	Sum of squares	Df	Mean Square	F	Sig.	Power
Schema_Type	10730,09	1	10730,09	0,01	0,95	0,10

Table 2. ANOVA results for undergraduate students from UCLM (Spain)

Source	Sum of squares	Df	Mean Square	F	Sig.	Power
Schema_Type	8557,55	1	8557,55	0,09	0,82	0,10

Table 3. ANOVA results for PhD students from Pinar del Rio University (Cuba)

Source	Sum of squares	Df	Mean Square	F	Sig.	Power
Schema_Type	101117,21	1	101117,21	0,02	0,90	0,10

As a conclusion we can deduce that there seems to be no difference in the understandability of the schematas because of the design method used. Nevertheless, as these conclusions could be due to the sizes of the schematas used in the experiment (they are not very large) or to the fact that SQL has been used to obtain the answers, we decided to replicate the experiment incorporating some changes that allowed us to avoid these limitations as far as possible.

4. Second experiment

In the second experiment and taking into account the previously obtained results, we decided to make a replica in which, working with the same hypotheses and with the same variables, we incorporated two fundamental changes; firstly we decided to increase the number of objects and, secondly, we tried to facilitate the work of the subjects allowing them to use natural language instead of the SQL.

Thus, the new experiment can be characterised as follows.

4.1. Experimental design

In this occasion, the experiment consisted of twelve schemes (the six ones from the first experiment and other six new schemas), six traditional schematas and six semantically equivalent schematas designed with star diagrams. On each one of them, the subjects had to indicate (in natural language), the necessary steps to obtain certain data from the datawarehouse schema and to write down the time (in seconds) that they took to perform these steps (as in the previous case, since, as we have already indicated, the dependent variable did not vary).

4.2. Subjects

In this occasion we had eighteen people, final year undergraduate students of the School of Computer Science of Ciudad Real (UCLM) who were doing a course on Information Retrieval where all the concepts related to datawarehouses were explained. In addition, all of them had attended a course on Databases (mandatory at the third level of their studies) in which all the contents relative to the relational model are dealt with in depth.

4.3. Limitations

The main limitation in this case is the low number of subjects.

4.4. Results

We fixed a value $\alpha = 0.1$ and, as the design is the same as in the previous experiment, we applied also the repeated measure univariate ANOVA test (SPSS 11, 2001).

In table 4 the results obtained from the experiment are shown. As the value obtained is less than α, we can reject the null hypothesis, and therefore there is a difference in the time needed to answer the questions according to the type of design used (traditional or star, Schema_type variable).

Table 4. ANOVA results for the second experimental work

Source	Sum of squares	Df	Mean Square	F	Sig.	Power
Schema_Type	154615,005	1	154615,005	11,572	0,001	0,960

As the obtained value is significant, the next step is to take the Difference of means to obtain more data. In table 5 the results obtained for this statistical test appear.

Table 5. Results of the Difference of means for the second experimental work

Number of Schema	Traditional design	Star design	Difference of means
1	174,33	164,56	9,78
2	263,78	259,72	4,06
3	245,22	243,61	1,61
4	333,83	152,67	181,17
5	168,00	158,61	9,39
6	299,83	184,78	115,06
Total	247,50	193,99	53,51

Based on the results obtained for the Difference of means, it can be concluded that when the datawarehouse is designed using star diagrams the time averages are smaller than when we use traditional design, and so we could conclude that the star design seems to be easier to understand than the traditional one for the design of datawarehouses.

5. Conclusions from all the experimental work developed

As can be appreciated, the results obtained are not definitive. Although in the first study we obtained the result that both modelling techniques could be appropriate for the logical design of the datawarehouses in the second study the star model seems to be more suited.

So, it seems that, at least, the use of the star design is not more difficult than the traditional design.

Nevertheless, to be able to reach more definitive and trustworthy results it is essential to carry out replicas of the second experimental work with more subjects with differing experience (for example with professional designers of datawarehouses). In addition, it would be advisable to perform another type of replica, for example, varying the hypotheses.

6. Conclusions and future works

Datawarehouses are one of the main trends in information systems since they help in strategic decision making.

Diverse methods for datawarehouse design have been set out based on star diagrams, since it is supposed that these diagrams, compared with the use of traditional modeling (ER and relational), increase the effectiveness and the understanding of the schemes of datawarehouses.

Although this affirmation is widely accepted, it has not been empirically demonstrated, which is why we decided to do a series of experiments. The experiments try to detect causal relationships between the logical design of a datawarehouse (traditional vs star) and the understandability of it. In all of them subjects had to make queries (in SQL or natural language) about a logical datawarehouse schema (traditional or star design). The way to determine the understandability of each of the schematas was to record the time required to carry out the indicated operations.

As a conclusion of our study we can say that using the star model, as was anticipated, we obtain schematas not more difficult to understand than the relational ones and, in some cases, they have turned out to be simpler.

To be able to have reliable results and definitive conclusions it is necessary to perform more replicas of these experiments and also new experiments together with their replicas. In addition, it is also fundamental to run case studies to know if the results obtained from controlled experiments are the same.

ACKNOWLEDGMENT

This research is part of the MESSENGER project (PCI-03-001) supported by the Consejeria de Ciencia y Tecnología of Junta de Comunidades de Castilla-La Mancha (Spain) and the CALIPO project (TIC 2003-07804-C05-03) supported by the Ministerio de Ciencia y Tecnologia (Spain).

REFERENCES

[1] Hörst, M., B. Regnell and C. Wohlin (2000) Using students as Subjects – A Comparative Study of Students & Profesionals in Lead-Time Impact Assessment. *4th Conference on Empirical Assessment & Evaluation in Software Engineering, EASE,* Keele University, UK.

[2] Chenoweth, T., Schuff, D. and St.Louis, R. (2003). A method for developing dimensional data marts. Communications of the ACM. December 2003. Vol. 46, No.12. pp. 93-98

[3] Inmon, W.H. (2002) Building the data warehouse. 3rd ed. Ed. Wiley

[4] Jarke, M., Lenzerini, M., Vassilou, Y. and Vassiliadis, P. (2000) Fundamentals of Data Warehouses. Ed. Springer.

[5] Kimball, R., Reeves, L., Ross, M. and Thornthwaite, W. (1998) The datawarehouse life cycle toolkit. Ed. Wiley

[6] SPSS 11.0. (2001) *Syntax Reference Guide*,. Chicago, SPSS Inc., 2001.