



Association for  
Computing Machinery

*Advancing Computing as a Science & Profession*

Kaiserslautern, Germany  
October 9-10, 2008



# ESEM'08

Proceedings of the 2008 ACM-IEEE International Symposium on  
**Empirical Software Engineering and Measurement**

*Sponsored by:*

**ACM SIGSOFT & IEEE CS**

*Supported by:*

**Siemens AG & Robert Bosch GmbH**

Kaiserslautern, Germany  
October 9-10, 2008



Association for  
Computing Machinery

*Advancing Computing as a Science & Profession*



8

ACM-IEEE International Symposium on  
**Software Engineering and Measurement**

CS

**Bosch GmbH**



**Association for  
Computing Machinery**

*Advancing Computing as a Science & Profession*

**The Association for Computing Machinery**  
**2 Penn Plaza, Suite 701**  
**New York, New York 10121-0701**

Copyright © 2008 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: Publications Dept., ACM, Inc. Fax +1 (212) 869-0481 or <permissions@acm.org>.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

**Notice to Past Authors of ACM-Published Articles**

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that has been previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform [permissions@acm.org](mailto:permissions@acm.org), stating the title of the work, the author(s), and where and when published.

**ISBN: 978-1-59593-971-5**

Additional copies may be ordered prepaid from:

**ACM Order Department**  
PO Box 11405  
New York, NY 10286-1405

Phone: 1-800-342-6626  
(US and Canada)  
+1-212-626-0500  
(all other countries)  
Fax: +1-212-944-1318  
E-mail: [acmhelp@acm.org](mailto:acmhelp@acm.org)

**ACM Order Number 594087**

Printed in the USA

# Message from the General Chair

It is my great pleasure to welcome you to the Second International Symposium on Empirical Software Engineering and Measurement (ESEM 2008) in Kaiserslautern, Germany. Empirical software engineering and measurement are important sub-disciplines aimed at advancing software engineering towards an engineering discipline. They focus on measurement in order to better understand, control, and improve software development processes, and on empirical studies to better model products and processes and understand process products.

The 1<sup>st</sup> ESEM was held in Madrid in September 2007. Although this is only the second ESEM symposium, this series continues the long tradition of two high-quality conference series. ESEM resulted from the merger of the IEEE/ACM International Symposium on Empirical Software Engineering (ISESE) and the IEEE International Software Metrics Symposium (METRICS). ISESE was held five times, first in Japan in 2002 and last in Brazil in 2006. As its name suggests, ISESE's primary goal was to disseminate knowledge on and from experimental research on software engineering. METRICS is a long-standing conference series, which was held first in the United States in 1993 and last in Italy in 2005. METRICS' foremost objective was to disseminate knowledge on and from research on software metrics.

ESEM 2008 is part of the week-long Experimental Software Engineering International Week (ESEIWI) in Kaiserslautern, Germany. It is an honor to welcome this conference to Kaiserslautern. Kaiserslautern has evolved into one of the premier centers for software & systems research in Europe. About 800 scientists and engineers – spread across 3 departments of the University of Kaiserslautern and several institutes – are working on advancing the state of the art in our field. The institutes include a Max-Planck Institute on Software Systems, two Fraunhofer Institutes – for Industrial Mathematics (ITWM) and Experimental Software Engineering (IESE) – and the German Research Institute for Artificial Intelligence (DFKI). I invite you to learn more about any of these institutes during your stay.

We have a full two-day program of three parallel tracks on each day including two keynotes, technical papers, short papers, and a panel (please see the message from the Program Chairs for details).

Putting together *ESEM 2008* has been a team effort. First of all, I would like to thank the authors of all submitted papers. Furthermore, I would like to express my gratitude to the program committee and to all external reviewers, who worked very hard on reviewing papers and providing suggestions for their improvements. I would also like to thank Sebastian Elbaum and Jürgen Münch, the program co-chairs Maria Teresa Baldassarre - the short papers chair, and Haruka Nakao - the poster program chair, as well as Andreas Jedlitschka, the treasurer, and Marcus Ciolkowski, the local arrangements chair. Finally, I would like to thank our society sponsor, ACM SIGSOFT, for their continued support of these successful meetings, and our industrial supporters Robert Bosch GmbH and Siemens AG for their financial support for the social events surrounding the conference.

I hope that you will find this program interesting and thought-provoking and that the symposium will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

**Dieter Rombach**

*ESEM08 General Chair*

*University of Kaiserslautern & Fraunhofer IESE  
Kaiserslautern, Germany*

## Message from the Program Co-Chairs

We would like to welcome you to Kaiserlautern, Germany, for the Second International Symposium on Empirical Software Engineering and Measurement – ESEM 2008. This year's technical program continues to advance the field of empirical software engineering and metrics with thoroughly reviewed technical papers complemented by two remarkable keynotes and a vibrant panel.

The call for research papers attracted over 100 submissions. To assess these submissions, the program committee members performed over 320 reviews and carried out dozens of online discussions. This process resulted in the selection of 28 papers that form our technical program. The papers cover a variety of topics, ranging from observational studies of programmers to measurement mechanisms for assessing technologies and organizations. In addition to the technical papers, short paper sessions and a poster exhibition will help to provide further coverage of ongoing work.

The keynote speakers and the panel will broaden and enrich the program by providing insights and reflections on our community strengths, weaknesses, and promising research directions. The keynotes will be given by Professor Mary Shaw from Carnegie Mellon University, USA, and Mr. Harald Hönninger, vice-president of Bosch corporate research division. The panel consists of highly renowned experts and promises interesting insights with respect to success stories, lessons learned from failures, and the next frontiers of empirical software engineering.

We would like to thank the authors who provided the content for the program, and to express our sincere gratitude to the program committee members and the external reviewers who volunteered time and resources to make this program possible. We would also like to thank the General Chair, Dieter Rombach, for his support and guidance, and Andreas Jedlitschka and Marcus Ciolkowski for the local organization of this conference and the maintenance of the website.

We hope that you will enjoy this program and that the symposium provides you with a valuable opportunity to share ideas with other researchers and practitioners.

### **Sebastian Elbaum**

*Program Co-Chair  
University of Nebraska-Lincoln (USA)*

### **Jürgen Münch**

*Program Co-Chair  
Fraunhofer Institute for Experimental  
Software Engineering (Germany)*

# Table of Contents

ESEM'08 Organization.....x

## Keynote Address

Session Chair: J. Münch (*Fraunhofer Institute for Experimental Software Engineering*)

- **Using Empirical Methods to Improve Industrial Technology Transfer**.....1  
Harald Hoeningner, Mark Müller (*Robert Bosch GmbH*)

## Full Papers

### Session 1A: Coordination and Communication

Session Chair: F. Lanubile (*University of Bari*)

- **Socio-Technical Congruence: A Framework for Assessing the Impact of Technical and Work Dependencies on Software Development Productivity**.....2  
Marcelo Cataldo (*Bosch Corporate Research*), James D. Herbsleb, Kathleen M. Carley (*Carnegie Mellon University*)
- **A Multiple Case Study Investigating the Interaction between Manufacturing and Development Organizations in Automotive Software Engineering**.....12  
Joakim Pernstål, Ana Magazinic, Peter Öhman (*Chalmers, Computer Science & Engineering*)

### Session 1B: Testing and Analysis

Session Chair: S. Vegas (*Universidad Politécnica de Madrid*)

- **Empirical Evaluations of Regression Test Selection Techniques: A Systematic Review**..22  
Emelie Engström, Mats Skoglund, Per Runeson (*Lund University*)
- **Capture-recapture in Software Unit Testing — A Case Study**.....32  
Hanna Scott, Claes Wohlin (*Blekinge Institute of Technology*)
- **On Establishing a Benchmark for Evaluating Static Analysis Alert Prioritization and Classification Techniques**.....41  
Sarah Heckman, Laurie Williams (*North Carolina State University*)

### Session2A: Estimation Models I

Session Chair: C. Seaman (*University of Maryland, Baltimore*)

- **Comparative Studies of the Model Evaluation Criteria MMRE and PRED in Software Cost Estimation Research**.....51  
Dan Port (*University of Hawaii at Manoa*), Marcel Korte (*University of Applied Sciences and Arts Dortmund*)
- **Phase Distribution of Software Development Effort**.....61  
Ye Yang, Mei He, Mingshu Li, Qing Wang (*Chinese Academy of Sciences*), Barry Boehm (*University of Southern California*)
- **Combining Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation**.....70  
Nikolaos Mitrás, Lefteris Angelis (*Aristotle University of Thessaloniki*)

### Session2B: Modeling and Architecture

Session Chair: J. Carver (*University of Alabama*)

- **An Industrial Case Study of Architecture Conformance**.....80  
Jacek Rosik (*Lero, University of Limerick*), Andrew Le Gear, Jim Buckley (*University of Limerick*), Muhammad Ali Babar (*Lero, University of Limerick*)
- **A Survey into the Rigor of UML Use and its Perceived Impact on Quality and Productivity**.....90  
Ariadi Nugroho, Michel R.V. Chaudron (*Leiden University*)
- **Model-based Functional Size Measurement**.....100  
Luigi A. Lavazza, Vicri del Bianco (*University of Insubria*), Carla Garavaglia (*Syrea - Intecy*)

## Keynote Address

Session Chair: S. Elbaum (*University of Nebraska*)

- **Empirical Challenges in Ultra Large Scale Systems** ..... 110  
Mary Shaw (*Carnegie Mellon University*)

## Session 3A: From the Programmers' Trenches

Session Chair: H. Erdogmus (*NRC Institute for Information Technology*)

- **Problems in Agile Trenches** ..... 111  
Mira Kajko-Mattsson (*Stockholm University and Royal Institute of Technology*)
- **Pair Programming: What's in it for Me?** ..... 120  
Andrew Begel, Nachiappan Nagappan (*Microsoft Research*)
- **Why Do Programmers Avoid Metrics?** ..... 129  
Medha Umarji, Carolyn Seaman (*University of Maryland, Baltimore County*)

## Session 3B: Inspections

Session Chair: D. Pfahl (*Simula Research Laboratory*)

- **The Impact of Time Controlled Reading on Software Inspection Effectiveness and Efficiency: A Controlled Experiment** ..... 139  
Kai Petersen, Kari Rönkkö, Claes Wohlin (*Blekinge Institute of Technology*)
- **Defect Categorization: Making Use of a Decade of Widely Varying Historical Data** ..... 149  
Carolyn B. Seaman, Forrest Shull, Myrna Regardie, Denis Elbert, Raimund L. Feldmann (*University of Maryland*),  
Yuepu Guo (*University of Maryland Baltimore County*),  
Sally Godfrey (*NASA Goddard Space Flight Center*)
- **Evaluation of Capture-Recapture Models for Estimating the Abundance of Naturally-Occurring Defects** ..... 158  
Gursimran Singh Walia (*Mississippi State University*), Jeffrey C Carver (*University of Alabama*)

## Session 4A: Metrics and Methodology

Session Chair: M. Oivo (*University of Oulu*)

- **Some Lessons Learned in Conducting Software Engineering Surveys in China** ..... 168  
Junzhong Ji (*Beijing University of Technology*),  
Jingyue Li, Reidar Conradi (*Norwegian University of Science and Technology*),  
Chunhuan Liu (*Beijing University of Technology*),  
Jianqiang Ma (*Norwegian University of Science and Technology*),  
Weibing Chen (*Beijing University of Technology*)
- **Strength of Evidence in Systematic Reviews in Software Engineering** ..... 178  
Tore Dybå, Torgeir Dingsøyir (*SINTEF ICT*)
- **Refining the Axiomatic Definition of Internal Software Attributes** ..... 188  
Sandro Morasca (*Università degli Studi dell'Insubria*)

## Session 4B: Faults and Failures

Session Chair: L. Williams (*North Carolina State University*)

- **Quantitative Analysis of Faults and Failures with Multiple Releases of SoftPM** ..... 198  
Shujian Wu, Qing Wang, Ye Yang (*The Chinese Academy of Sciences*)
- **Iterative Identification of Fault-Prone Binaries Using In-Process Metrics** ..... 206  
Lucas Layman (*North Carolina State University*),  
Gunnar Kudrjavets, Nachiappan Nagappan (*Microsoft Corporation*)

## Session 5A: Estimation Models II

Session Chair: M. Ciolkowski (*Fraunhofer Institute for Experimental Software Engineering*)

- **A Constrained Regression Technique for COCOMO Calibration** ..... 213  
Vu Nguyen, Bert Steece, Barry Boehm (*University of Southern California*)
- **Reducing Biases in Individual Software Effort Estimations: A Combining Approach** ..... 223  
Qi Li, Qing Wang, Ye Yang, Mingshu Li (*Chinese Academy of Sciences*)
- **Any Other Cost Estimation Inhibitors?** ..... 233  
Ana Magazinic, Joakim Pernstål (*Chalmers*)
- **Session 5B: From the Manager's Trenches**  
Session Chair: P. Runeson (*Lund University*)
- **Empirical Results from Using Custom-Made Software Project Control Centers in Industrial Environments** ..... 243  
Marcus Ciolkowski, Jens Heidrich (*Fraunhofer IESE*), Frank Simon (*SQS AG*), Mathias Radtke (*BTU Cottbus*)
- **A Survey on Software Cost Estimation in the Chinese Software Industry** ..... 253  
Da Yang, Qing Wang, Mingshu Li, Ye Yang, Kai Ye, Jing Du (*Chinese Academy of Sciences*)
- **A Survey of Software Project Managers on Software Process Change** ..... 263  
Yuepu Guo, Carolyn B. Seaman (*University of Maryland, Baltimore County*)

## Short Papers

### Session 1C: Evaluation and Comparison of Techniques and Models

Session Chair: M. Genero (*University of Castilla La Mancha*)

- **A Pilot Study of Comparative Customer Comprehension between Extreme X-Machine and UML Models** ..... 270  
Christopher Thomson, Mike Holcome, Tony Cowling, Tony Simons, George Michaelides (*University of Sheffield*)
- **Enhancing Predictive Models Using Principal Component Analysis and Search Based Metric Selection: A Comparative Study** ..... 273  
Rodrigo Vivanco, Dean Jin (*University of Manitoba*)
- **Evaluating the Usefulness of Software Visualization in Supporting Software Comprehension Activities** ..... 276  
Glauco de F. Carneiro, Rodrigo Magnavita, Eduardo Spinola, Fábio Spinola, Manoel Mendonça (*Salvador University*)
- **A Hybrid Faulty Module Prediction Using Association Rule Mining and Logistic Regression Analysis** ..... 279  
Yasutaka Kamei, Akito Monden, Shuji Morisaki, Ken-ichi Matsumoto (*Nara Institute of Science and Technology*)
- **Mining Software Code Repositories and Bug Databases using Survival Analysis Models** ..... 282  
Michael Wedel (*Universität Stuttgart*), Uwe Jensen (*University of Hohenheim*), Peter Göhner (*Universität Stuttgart*)
- **Adding Planned Design to XP Might Help Novices' Productivity (or Might Not): Two Controlled Experiments** ..... 285  
René Noël, Gonzalo Valdes, Marcello Visconti, Hernán Astudillo (*Universidad Técnica Federico Santa María*)



## Session 2C: Empirical Studies of Processes and Products

Session Chair: D. Winkler (*Vienna University of Technology*)

- **Using Students as Subjects — An Empirical Evaluation** ..... 288  
Mikael Svahnberg (*Blekinge Institute of Technology*), Aybülke Aarum (*University of New South Wales*), Claes Wohlin (*Blekinge Institute of Technology*)
- **Empirical Study of How Personality, Team Processes and Task Characteristics Relate to Satisfaction and Software Quality** ..... 291  
Silvia T. Acuña (*Universidad Autónoma de Madrid*), Marta N. Gómez (*Universidad San Pablo-CEU*), Juan de Lara (*Universidad Autónoma de Madrid*)
- **Empirical Evaluation of Analogy-X for Software Cost Estimation** ..... 294  
Jacky Keung (*NICTA Ltd. and University of New South Wales*)
- **Improving Application and Understanding of Experience Packages through Learning Spaces** ..... 297  
Eric Ras (*Fraunhofer IESE*)
- **Does the Use of Stereotypes Improve the Comprehension of UML Sequence Diagrams?** ..... 300  
Marcela Genero, José A. Cruz-Lemus (*University of Castilla-La Mancha*), Danilo Caivano (*University of Bari*), Sílvia Abrahão, Emilio Infrán, José A. Carsi (*Universidad Politécnica de Valencia*)
- **Web Application Fault Classification — An Exploratory Study** ..... 303  
Yuepu Guo, Sreedevi Sampath (*University of Maryland, Baltimore County*)

## Session 3C: Development of Predictive Models

Session Chair: O. Dieste (*Universidad Politécnica de Madrid*)

- **Exposure Model for Prediction of Number of Customer Reported Defects** ..... 306  
Keld Raaschou (*SimCorp A/S*), Austen W. Rainer (*University of Hertfordshire*)
- **Analysis of the Reliability of a Subset of Change Metrics for Defect Prediction** ..... 309  
Raimund Moser (*Free University of Bolzano-Bozen*), Witold Pedrycz (*University of Alberta*), Giancarlo Succi (*Free University of Bolzano-Bozen*)
- **An Over-sampling Method for Analogy-based Software Effort Estimation** ..... 312  
Yasutaka Kamei (*Nara Institute of Science and Technology*), Jacky Keung (*National ICT Australia Ltd.*), Akito Monden, Ken-ichi Matsumoto (*Nara Institute of Science and Technology*)
- **An Empirical Model to Predict Security Vulnerabilities using Code Complexity Metrics** ..... 315  
Yonghee Shin, Laurie Williams (*North Carolina State University*)
- **Ensemble of Software Defect Predictors: A Case Study** ..... 318  
Ayse Tosun, Burak Turhan, Ayse Bener (*Bogazici University*)
- **Managing Software Quality through a Hybrid Defect Content and Effectiveness Model** ..... 321  
Michael Kläs, Frank Elberzhager (*Fraunhofer Institute for Experimental Software Engineering*), Haruka Nakao (*Japan Manned Space Systems Corporation*)

## Session 4C: Experience in Process Improvement

Session Chair: T. Gorschek (*BTH*)

- **Surveying Model Based Testing Approaches Characterization Attributes** ..... 324  
Arlito Claudio Dias Neto, Guilherme Horta Travassos (*Federal University of Rio de Janeiro*)
- **Statistical Process Control for Software: A Systematic Approach** ..... 327  
Nicola Boffoli, Giovanni Bruno, Danilo Caivano, Gemma Mastelloni (*University of Bari*)
- **Issues and Effort in Integrating Data from Heterogeneous Software Repositories and Corporate Databases** ..... 330  
Rudolf Ramler, Klaus Wolfmeier (*Software Competence Center Hagenberg*)

- **A Defect-Driven Process for Software Quality Improvement**.....333  
Brian Robinson, Patrick Francis, Fredrik Ekdahl (*ABB Robotics*)
- **IMPS: An Experimentation Based Investigation of a Nationwide Software Development Reference Model**.....336  
Marcos Kalinowski (*COPPE/UFRJ*), Kival Chaves Weber (*SOFTEX*), Guilherme Horta Travassos (*COPPE/UFRJ*)
- **Using the ProdFLOW™ Approach to Address the Myth of Productivity in R&D Organizations** .....339  
Melanie Rube (*Siemens AG, CT SE3*), Stefan Wagner (*Technical University of Munich*)

**Session 5C: Empirical Evidence and Systematic Review**

Session Chair: D. Caivano (*University of Bari*)

- **A Mapping Study on Empirical Evidence related to the Models and Forms used in the UML** .....342  
Rialette Pretorius, David Budgen (*Durham University*)
- **Software Process Simulation over the Past Decade: Trends Discovery from A Systematic Review** .....345  
He Zhang (*National ICT Australia*), Barbara Kitchenham (*Keele University*), Dietmar Pfahl (*University of Oslo*)
- **An Empirical Investigation of Scenarios Gained and Lost in Architecture Evaluation Meetings**.....348  
Dietmar Winkler, Stefan Biffl (*Vienna University of Technology*), Muhammad Ali Babar (*Lero, University of Limerick*)
- **Are Good Code Reviewers Also Good at Design Review?**.....351  
Hideake Uwano, Akito Monden, Ken-ichi Matsumoto (*Nara Institute of Science and Technology*)
- **Understandability Measurement in an Early Usability Evaluation for Model-Driven Development: An Empirical Study** .....354  
Jose Ignacio Panach, Nelly Condori-Fernández, Francisco Valverde, Nathalie Aquino, Óscar Pastor (*Universidad Politécnica de Valencia*)

**Posters**

- **Automatic Extraction of the Main Terminology used in Empirical Software Engineering through Text Mining Techniques**.....357  
Francisco P. Romero, José A. Olivás, Marcela Genero, Mario Piattini (*University of Castilla La Mancha*)
- **Exploring Effort Distribution in RUP Projects** .....359  
Werner Heijstek, Michel R.V. Chaudron (*Leiden University*)
- **Fit Data Selection for Software Effort Estimation Models**.....360  
Koji Toda, Akito Monden, Ken-ichi Matsumoto (*NARA Institute of Science and Technology*)

- Organization**.....362

# ESEM 2008 Organization

## General Chair:

Dieter Rombach - University of Kaiserslautern and Fraunhofer Institute for Experimental Software Engineering (Germany)

## Program Chairs:

Sebastian Elbaum - University of Nebraska-Lincoln (USA)  
Jürgen Münch - Fraunhofer Institute for Experimental Software Engineering (Germany)

## Steering Committee:

Natalia Juristo - Universidad Politécnica de Madrid (Spain)  
Sebastian Elbaum - University of Nebraska - Lincoln (USA)  
James Miller - University of Alberta (Canada)  
Jürgen Münch - Fraunhofer IESE (Germany)  
Dieter Rombach - University of Kaiserslautern and Fraunhofer IESE (Germany)  
Carolyn Seaman - University of Maryland at Baltimore County (USA)  
Rick Selby - Northrop Grumman (USA)  
Sira Vegas - Universidad Politécnica de Madrid (Spain)  
Laurie Williams - North Carolina State University (USA)

## Program Committee:

Silvia Abraham - Universidad Politecnica de Valencia (Spain)  
Muhammad Ali Barbar - Lero, University of Limerick (Ireland)  
Anneliese Amschler Andrews - University of Denver (USA)  
James Andrews - University of Western Ontario (Canada)  
Stefan Biffli - TU Vienna (Austria)  
Lionel Briand - Simula Research Laboratory (Norway)  
Luigi Buglione - Engineering.IT (Italy)  
Gerardo Canfora - University of Sannio (Italy)  
Jeff Carver - University of Alabama (USA)  
Marcus Ciolkowski - University of Kaiserslautern (Germany)  
Reidar Conradi - Norwegian University of Science and Technology (Norway)  
Tore Dyba - Sintef (Norway)  
Christof Ebert - Alcatel (France)  
Khaled El Emam - University of Ottawa (Canada)  
Hakan Erdogmus - NRC Institute for Information Technology (Canada)  
Marcela Genero - University of Castilla La Mancha (Spain)  
Tracy Hall - University of Hertfordshire (UK)  
Lorin Hochstein - University of Nebraska-Lincoln (USA)  
Martin Höst - Lund University (Sweden)  
Frank Houdek - Daimlerchrysler (Germany)  
Andreas Jedlitschka - Fraunhofer IESE (Germany)  
Ross Jeffery - University of New South Wales (Australia)  
Philip Johnson - University of Hawaii (USA)  
Magne Jorgensen - Simula Research Laboratory (Norway)  
Natalia Juristo - Universidad Politecnica de Madrid (Spain)  
Shinji Kusumoto - Osaka University (Japan)

**Program Committee**  
**(Continued):**

Filippo Lanubile - University of Bari (Italy)  
Jose Carlos Maldonado - Universidade de Sao Paulo (Brazil)  
Kenichi Matsumoto - Nara Institute of Science and Technology (Japan)  
Emilia Mendes - University of Auckland (New Zealand)  
James Miller - University of Alberta (Canada)  
Audris Mockus - Avaya Labs Research (USA)  
Sandro Morasca - University degli Studi dell'Insubria (Italy)  
Maurizio Morisio - Politecnico di Torino (Italy)  
Nachi Nagappan - Microsoft (USA)  
Markku Oivo - University of Oulu (Finland)  
Dietmar Pfahl - Simula Research Laboratory & University of Oslo (Norway)  
Mario Piattini - University of Castilla La Mancha (Spain)  
Marc Roper - University of Strathclyde (UK)  
Guenther Ruhe - University of Calgary (Canada)  
Per Runeson - Lund University (Sweden)  
Sreedevi Sampath - University of Maryland, Baltimore (USA)  
Carolyn Seaman - University of Maryland, Baltimore (USA)  
Rick Selby - Northrop Grumman (USA)  
Martin Shepperd - Brunel University (UK)  
Forrest Shull - Fraunhofer Center, Maryland (USA)  
Janice Singer - National Research Council (Canada)  
Harvey Siy - University of Nebraska (USA)  
Dag Sjoberg - Simula Research Laboratory (Norway)  
Arie Van Deursen - Delft University of Technology (Netherlands)  
Sira Vegas - Universidad Polit cnica de Madrid (Spain)  
June Verner - NICTA (Australia)  
Laurie Williams - North Carolina State University (USA)  
Claes Wohlin - Blekinge Institute of Technology (Sweden)

**Finance Chair:** Andreas Jedlitschka - Fraunhofer Institute for Experimental Software Engineering (Germany)

**Local Arrangements Chair:** Marcus Ciolkowski - Fraunhofer Institute for Experimental Software Engineering (Germany)

**Publicity Chair:** James Miller - University of Alberta (Canada)

**Proceedings Chair:** Marcus Ciolkowski - Fraunhofer Institute for Experimental Software Engineering (Germany)

**Short Papers Chair:** Teresa Baldassarre - Universita' degli Studi di Bari (Italy)

**Poster Chair:** Haruka Nakao - Japan Manned Space Systems Cooperation (Japan)

# Does the Use of Stereotypes Improve the Comprehension of UML Sequence Diagrams?

Marcela Genero,  
José A. Cruz-Lermus  
ALARCOS Research Group.  
Department of Technologies and  
Information Systems.  
University of Castilla-La Mancha  
Paseo de la Universidad, 4  
13071 Ciudad Real, Spain  
0034 926295300 ext. 3740/3748

{Marcela.Genero,  
JoseAntonio.Cruz}@uclm.es

Danilo Caivano  
Department of Informatics  
University of Bari  
Via E. Orabona, 4  
70126 Bari, Italy  
caivano@di.uniba.it

Silvia Abrahão,  
Emilio Infran,  
José Angel Carsí  
ISSI Research Group  
Universidad Politécnica de Valencia  
Camino de Vera s/n  
46022, Valencia, Spain  
{sabrahao, einsfran,  
pcarsj}@dsic.upv.es

## ABSTRACT

This paper reports on a controlled experiment that investigates the influence of stereotypes in UML sequence diagrams. The comprehension of UML sequence diagrams with and without stereotypes is analyzed from three different perspectives: semantic comprehension, retention and transfer. The experiment was carried out with 77 undergraduate students of Computer Science from the University of Bari in Italy. The results obtained show a slight tendency in favor of the use of stereotypes in facilitating the comprehension of UML sequence diagrams. Further replications are needed to obtain more conclusive results.

## Categories and Subject Descriptors

D.2.2 [Software Engineering]: Design tools and techniques - *object-oriented design methods*.

## General Terms

Design, Experimentation.

## Keywords

UML sequence diagrams, stereotypes, comprehension, understandability, controlled experiment.

## 1. INTRODUCTION

Stereotypes are used to clarify or extend the meaning of model elements. However, the most intuitive notations are not always the best performing ones. In recent years, several studies that investigate the use of stereotypes have been conducted in the field of Software Engineering [5, 6]. These studies are focused on investigating the influence of stereotypes on the comprehension of UML class and collaboration diagrams. It is widely recognized that comprehension is one of the main factors to influence maintainability, which must be assessed from the first steps of the software lifecycle.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.  
*ESEM'08*, October 9–10, 2008, Kaiserslautern, Germany.  
Copyright 2008 ACM 978-1-59593-971-5/08/10...\$5.00.

In this paper, we present a controlled experiment to assess whether the use of stereotypes improves the comprehension of UML sequence diagrams. Specifically, we consider a set of stereotypes that have been proposed to enrich the semantics of interaction messages in UML sequence diagrams in the context of a Requirements Engineering approach for model-driven software development [3].

The structure of the paper is as follows. The definition of stereotypes for UML sequence diagrams is introduced in Section 2. The description and execution of the experiment is described in Section 3. The data analysis and the interpretation of the experiment results are presented in Section 4. Finally, the conclusions and future research work are presented in Section 5.

## 2. THE STEREOTYPES

In UML, sequence diagrams are a means to model an aspect of the dynamic behavior of a system. They can be used in the context of the whole system, of a subsystem, or they can be attached to a use case. Some authors indicate that at least one sequence diagram should be drawn per use case.

These diagrams can be used to graphically represent a pattern interaction between objects by sending and receiving messages as time advances. The following set of UML stereotypes has been proposed [3] as a means to enhance the meaning of the object interactions (messages):

- **«signal»** this is a message between an actor and the system. The objective is to describe an interaction between an external entity and the system.
- **«service»** this is a message that updates the state of an object. The properties new/destroy/update can be used if the object is to be created, destroyed or modified.
- **«query»** this is a message to query the state of a specific object or set of objects.
- **«connect»** this is a message which is used to indicate that an object from the sender object type needs to be connected (or related) to an object from the receiver object type.

The goal of these stereotypes is twofold. First, to improve the comprehension of UML sequence diagrams. Second, to give

information to the transformation process of UML sequence diagrams into conceptual models following a model-driven development process.

### 3. EXPERIMENT DESCRIPTION

The goal of this experiment is to assess whether the use of stereotypes improves the understanding of UML sequence diagrams. Therefore, by using the GQM template for goal definition [1], the goal of our experiment is defined as follows: "Analyze the use of stereotypes for the purpose of evaluating it with respect to the comprehension of UML sequence diagrams from the point of view of the researcher, in the context of fourth year undergraduate students in Computer Science from the University of Bari".

To run and report this experiment we followed the recommendations provided in [7]. We selected a balanced factorial design with group-interaction confounding.

The experiment was carried out by 77 students who were taking a course in Empirical Software Engineering. The subjects had participated in two Software Engineering courses in which they had acquired training in UML diagrams. In addition, they were taught a three hour lecture about the stereotypes for UML sequence diagrams in a training session organized to take place the day before the experiment was carried out. To avoid social threats due to evaluation apprehension, the students were not graded on their performance. They were granted with 2 extra points in their final valuation at the end of the course.

The *experimental objects* consisted of four diagrams, summarized below (they complete materials can be found at <http://alarcos.esi.ucm.es/ExpStereotypes>):

- DA-S: stereotyped diagram A and a general description of each type of stereotype.
- DA-N: non-stereotyped diagram A.
- DB-S: stereotyped diagram B and a general description of each type of stereotype.
- DB-N: non-stereotyped diagram B.

There were two *independent variables* in the experiment, the diagram type, with values: S (stereotyped) and N (non-stereotyped), and the diagram domain (A and B). UML sequence diagram A-x describes a car rental domain whereas diagram B-x describes a hotel domain. By combining each level of the independent variables we obtain four treatments. These experimental objects were presented in Italian.

In accordance with certain suggestions concerning the measurement of comprehension [2], we have used the Cognitive Theory of Multimedia Learning (CTML) [4]. By following CTML, the comprehension of the UML sequence diagrams has been defined through three variables: semantic comprehension, retention and transfer. *Semantic comprehension* is the ability to comprehend the semantics of the models. *Retention* is related to the comprehension of material being presented, and the ability to retain knowledge from it. *Transfer* is the ability to use the knowledge gained from the material to solve related problems which are not directly answerable from it.

In order to measure these variables, we have used three separate tests based on questionnaires. Each comprehension measure is computed as:

- **Effectiveness:** the proportion of correct answers provided in each test (number of correct answers / number of questions). This measure reflects the ability to understand the presented material correctly.
- **Efficiency:** the proportion of correct answers divided by time (Effectiveness / Time).

In this work, we call these values SCEffec/SCEffic (Semantic Comprehension Effectiveness and Efficiency), TransEffec/TransEffic (Transfer Effectiveness and Efficiency), and finally, RetenEffec/RetenEffic (Retention Effectiveness and Efficiency). For Semantic Comprehension we wished to test the following *hypotheses*:

- $H_{1,1,0}$ : stereotypes do not improve the subjects' SCEffec when attempting to comprehend a UML sequence diagram.  
 $H_{1,1,1} = -H_{1,1,0}$
- $H_{1,2,0}$ : stereotypes do not improve the subjects' SCEffic when attempting to comprehend an UML sequence diagram.  
 $H_{1,2,1} = -H_{1,2,0}$

We analogously formulated a set of hypothesis  $H_3$  for Retention measures (RetenEffec and RetenEffic), and another set  $H_4$  related to the Transfer measures (TransEffec and TransEffic).

The experiment consisted of two rounds. In each round, each of the groups was given a different treatment. In order to alleviate learning effect, we divided the experimental subjects into 4 groups. We randomly assigned the corresponding diagrams to each group, but in different order.

### 4. ANALYSIS AND INTERPRETATION

In this section, we discuss, for each variable, the descriptive statistics and the results of the ANOVA tests carried out to test the formulated hypotheses. All the statistical analyses presented in this section were performed using SPSS, with  $\alpha = 0.05$ .

Tables 1-3 present the descriptive statistics for the all the variables studied: Semantic comprehension, Retention and Transfer.

Table 1. Descriptive statistics for Semantic Comprehension

Diag.	Dom.	N	SCEffec		SCEffic	
			Mean	Std. Dev.	Mean	Std. Dev.
S	A	38	0.756	0.1410	0.0020	0.00069
	B	39	0.782	0.1312	0.0027	0.00133
	Total	77	0.769	0.1360	0.0023	0.00112
N	A	39	0.761	0.1079	0.0022	0.00082
	B	38	0.741	0.1390	0.0024	0.00096
	Total	77	0.751	0.1242	0.0023	0.00089

For both the variables (SCEffec and SCEffic), the subjects obtained better results in the stereotyped diagrams. Obviously this means that the stereotyped diagrams were more comprehensible, less error prone and required less time to understand than non-stereotyped diagrams.

Table 2 shows that, related to Retention, the subjects obtained better results in effectiveness when they did not use stereotypes, while the scores were higher for efficiency when they worked with stereotyped diagrams. In addition, RetenEffic and RetenEffic values are significantly better for the diagrams of domain B. Thus, the subjects who used non-stereotyped diagrams

were able to recall a greater number of correct functionalities of the models, but the subjects with non-stereotyped diagrams performed better if we relate the number of correct answers with the time spent answering the test. This is probably due to the fact that stereotyped diagrams contain more information than those which are non-stereotyped and thus prove to be more difficult to remember (but not to understand!).

**Table 2. Descriptive statistics for Retention**

Diag.	Dom.	N	RetenEffic		RetenEffic	
			Mean	Std. Dev.	Mean	Std. Dev.
S	A	38	0.636	0.1579	0.636	0.0036
	B	39	0.784	0.1830	0.784	0.0019
	Total	77	0.709	0.1854	0.709	0.0029
N	A	39	0.637	0.1814	0.637	0.0011
	B	38	0.807	0.1408	0.807	0.0044
	Total	77	0.723	0.1823	0.723	0.0032

**Table 3. Descriptive statistics for Transfer**

Diag.	Dom.	N	TransEffic		TransEffic	
			Mean	Std. Dev.	Mean	Std. Dev.
S	A	38	0.682	0.2501	0.682	0.0011
	B	39	0.829	0.2404	0.829	0.0058
	Total	77	0.755	0.2547	0.755	0.0053
N	A	39	0.616	0.2248	0.616	0.0013
	B	38	0.821	0.3139	0.821	0.0064
	Total	77	0.719	0.2907	0.719	0.0059

The results for Transfer (Table 3) are the opposite of those for Retention. Better results were obtained for effectiveness when the subjects used the stereotyped diagrams, but they were worse for efficiency. Moreover, these results confirm that the stereotyped diagrams are more comprehensible and less error prone but that the quantity of information they include implies that more time is needed for diagram processing: in this case, and unlike the Semantic Comprehension, the students have to reconstruct the message flow in the sequence diagram and thus have to spend more time in understanding the semantic first and the syntactic of the diagrams later. This implies more time spent, which thus leads to lower efficiency.

We carried out ANOVA tests to check the hypotheses presented in Section 3.1. From these analyses, we observed that in most cases, the domain significantly affects the comprehension of the diagrams (e.g., RetenEffic = 0.000; TransEffic = 0.000; TransEffic = 0.000). The results indicated that the subjects performed worse with the diagrams of domain A (car rental). This could imply that the subjects found the car rental domain more difficult. With respect to the use of stereotypes, there are no significant values in the tests (SCEffic = 0.388; RetenEffic = 0.650; TransEffic = 0.373; SCEffic = 0.578; RetenEffic = 0.723; TransEffic = 0.693). Moreover, the powers of the tests are low (e.g., SCEffic = 0.138; TransEffic = 0.144). This implies that we cannot reject the null hypotheses investigated in this work and accepting the alternative hypotheses could be risky and misleading. Furthermore, we have results that might be a consequence of the treatment or be determined by an interaction effect between treatment and domain.

## 5. CONCLUSIONS AND FUTURE WORK

This paper has investigated the use of stereotypes in the context of UML sequence diagrams. This took place through a controlled experiment and, unfortunately the obtained results cannot be assumed to be final. The p-values observed in the hypotheses test are non-significant and thus the null-hypotheses cannot be rejected. This implies that the stereotypes do not significantly improve comprehension in a statistical sense. Nevertheless, in 4 out of 6 cases the scores obtained by the subjects were higher when using stereotypes, which could imply that using stereotypes does, to some extent, help in the diagram comprehension.

Obviously these conflicting elements stimulate further investigation which we wish to carry out in the near future. We shall attempt to control the threats in a more effective manner and improve the experimental schema, addressing the following issues: (a) the domain effect must be eliminated before conclusions with regard to domains with a similar level of difficulty are drawn; (b) we must confront the experience factor by attempting to point out the effect that the experience of using our stereotypes had on the subjects' performance; (c) data analysis must be improved by clarifying the difference between the non unanswered questions and the wrong answers and their impact on Effectiveness and Efficiency. In this work they were dealt with in the same way; (d) to investigate the possible existence of other relevant factors [5]. The present work is the first experiment within a planned family of experiments, and it considers not only students but also practitioners.

## 6. ACKNOWLEDGEMENTS

This research is part of the META project (TIN2006-15175-C05-05), the IDONEO project (PAC08-0160-6141), and the Quality-driven Model Transformations project (JPV). The authors thank professors Giuseppe Visaggio and Maria Teresa Baldassarre and their students for having cooperated with this experiment.

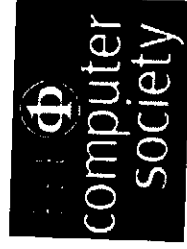
## 7. REFERENCES

- [1] Basili, V. and Rombach, H. 1998. The TAME Project: Towards Improvement-oriented Software Environments. *IEEE Trans. on Soft. Eng.* 14 (6), 758-773.
- [2] Gemino, A. and Wand, Y. 2005. Complexity and Clarity in Conceptual Modeling: Comparison of Mandatory and Optional Properties. *Data Know. Eng.* 55, 301-326.
- [3] Insfran, E., Pastor, P., and Wieringa, R. 2002. Requirements Engineering-based Conceptual Modeling. *J. Req. Eng.* 7(2), 61-72.
- [4] Mayer, R.E. 2001. *Multimedia Learning*. Cambridge Press.
- [5] Ricca, F., Di Penta, M., Torchiano, M., Tonella, P., and Ceccato, M. 2007. The Role of Experience and Ability in Comprehension Tasks supported by UML Stereotypes. In *Proc. of ICSE 2007* (Minneapolis, USA), 375-384.
- [6] Staron, M., Kuzniarz, L., and Wohlin, C. 2006. Empirical Assessment of Using Stereotypes to Improve Comprehension of UML Models: a Set of Experiments. *J. Syst.Soft.* 79, 727-742.
- [7] Wohlin, C., Runeson, P., Hast, M., Ohlsson, M.C., Regnell, B., and Wesslen, A. 2000. Experimentation in Software Engineering: an Introduction, Kluwer Academic Pub.

**Additional reviewers:**

Martin Auer	Lucas Layman
Maria Teresa Baldassarre	Otávio Lemos
Fabio Calefato	Jingyue Li
Marcio Delamaro	Teresa Mallardo
Oscar Dieste	Rajwinder Kaur Panesar-Walawege
Evgenia Egorova	Francisco Pino
Mohamed El-Attar	Daniel N. Port
Fabiano Ferrari	Yonghee Shin
Michael Gegick	Ben Smith
Domenico Gendarmi	Martin Soto
Stein Grimstad	Tor Stålhan
Anna Grimán	Marco Torchiano
Lauren Hayward	Adam Trendowicz
Sarah Smith Heckman	Irfan Ullah
Jens Heidrich	Giuseppe Visaggio
Toan Huynh	Gursimran Walia
Vigdis By Kampenes	Byron Williams
Jacky Keung	

**Sponsors:**



**Supporters:**

**SIEMENS**

Invented for life



**BOSCH**



## Author Index

- |                                |          |                            |                    |                            |               |
|--------------------------------|----------|----------------------------|--------------------|----------------------------|---------------|
| Abrahão, Silvia .....          | 300      | Göhner, Peter .....        | 282                | Nakao, Haruka .....        | 321           |
| Acuña, Silvia T. ....          | 291      | Gómez, Marta N. ....       | 291                | Nguyen, Vu .....           | 213           |
| Ali Babar, Muhammad .....      | 80       | Guo, Yuepu .....           | 149, 263, 303      | Noël, René .....           | 285           |
| Angelis, Lefteris .....        | 70       | He, Mei .....              | 61                 | Nugroho, Ariadi .....      | 90            |
| Aquino, Nathalie .....         | 354      | Heckman, Sarah .....       | 41                 | Öhman, Peter .....         | 12            |
| Astudillo, Hernán .....        | 285      | Heidrich, Jens .....       | 243                | Olivas, José A. ....       | 357           |
| Aurum, Aybüke .....            | 288      | Hejstek, Werner .....      | 359                | Panach, Jose Ignacio ..... | 354           |
| Babar, Muhammad Ali .....      | 348      | Herbsleb, James D. ....    | 2                  | Pastor, Óscar .....        | 354           |
| Begel, Andrew .....            | 120      | Hoeningger, Harald .....   | 1                  | Pedrycz, Witold .....      | 309           |
| Bener, Ayse .....              | 318      | Holcome, Mike .....        | 270                | Pernstål, Joakim .....     | 12, 233       |
| Biffi, Stefan .....            | 348      | Insfrán, Emilio .....      | 300                | Petersen, Kai .....        | 139           |
| Boehm, Barry .....             | 61, 213  | Jensen, Uwe .....          | 282                | Pfahl, Dietmar .....       | 345           |
| Boffoli, Nicola .....          | 327      | Ji, Junzhong .....         | 168                | Piattini, Mario .....      | 357           |
| Bruno, Giovanni .....          | 327      | Jin, Dean .....            | 273                | Port, Dan .....            | 51            |
| Buckley, Jim .....             | 80       | Kajko-Mattsson, Mira ..... | 111                | Pretorius, Rialette .....  | 342           |
| Budgen, David .....            | 342      | Kalinowski, Marcos .....   | 336                | Raaschou, Keld .....       | 306           |
| Caivano, Danilo .....          | 300, 327 | Kamei, Yasutaka .....      | 279, 312           | Radicke, Mathias .....     | 243           |
| Carley, Kathleen M. ....       | 2        | Keung, Jacky .....         | 294, 312           | Rainer, Austen W. ....     | 306           |
| Carneiro, Glauco de F. ....    | 276      | Kitchenham, Barbara .....  | 345                | Ramler, Rudolf .....       | 330           |
| Carsi, José A. ....            | 300      | Klås, Michael .....        | 321                | Ras, Eric .....            | 297           |
| Carver, Jeffrey C. ....        | 158      | Korte, Marcel .....        | 51                 | Regardie, Myrna .....      | 149           |
| Cataldo, Marcelo .....         | 2        | Kudrjavets, Gunnar .....   | 206                | Robinson, Brian .....      | 333           |
| Chaudron, Michel R. V. ....    | 90, 359  | Lavazza, Luigi A. ....     | 100                | Romero, Francisco P. ....  | 357           |
| Chen, Weibing .....            | 168      | Layman, Lucas .....        | 206                | Rönkkö, Kari .....         | 139           |
| Ciolkowski, Marcus .....       | 243      | Le Gear, Andrew .....      | 80                 | Rosik, Jacek .....         | 80            |
| Condori-Fernández, Nelly ..... | 354      | Li, Jingyue .....          | 168                | Ruhe, Melanie .....        | 339           |
| Conradi, Reidar .....          | 168      | Li, Mingshu .....          | 61, 223, 253       | Runeson, Per .....         | 22            |
| Cowling, Tony .....            | 270      | Li, Qi .....               | 223                | Sampath, Sreedevi .....    | 303           |
| Cruz-Lemus, José A. ....       | 300      | Liu, Chunnian .....        | 168                | Scott, Hanna .....         | 32            |
| de Lara, Juan .....            | 291      | Ma, Jianqiang .....        | 168                | Seaman, Carolyn .....      | 129, 149, 263 |
| del Bianco, Vieri .....        | 100      | Magazinovic, Ana .....     | 12, 233            | Shaw, Mary .....           | 110           |
| Dias Neto, Arilo Claudio ..... | 324      | Magnavita, Rodrigo .....   | 276                | Shin, Yonghee .....        | 315           |
| Dingsøyr, Torgeir .....        | 178      | Mastelloni, Gemma .....    | 327                | Shull, Forrest .....       | 149           |
| Du, Jing .....                 | 253      | Matsumoto, Ken-ichi .....  | 279, 312, 351, 360 | Simon, Frank .....         | 243           |
| Dybå, Tore .....               | 178      | Mendonça, Manoel .....     | 276                | Simons, Tony .....         | 270           |
| Ekdahl, Fredrik .....          | 333      | Michaelides, George .....  | 270                | Skoglund, Mats .....       | 22            |
| Elbert, Denis .....            | 149      | Mittas, Nikolaos .....     | 70                 | Spinola, Eduardo .....     | 276           |
| Elberzhager, Frank .....       | 321      | Monden, Akito .....        | 279, 312, 351, 360 | Spinola, Fábio .....       | 276           |
| Engström, Emelie .....         | 22       | Morasca, Sandro .....      | 188                | Steece, Bert .....         | 213           |
| Feldmann, Raimund L. ....      | 149      | Morisaki, Shuji .....      | 279                | Succi, Giancarlo .....     | 309           |
| Francis, Patrick .....         | 333      | Moser, Raimund .....       | 309                | Svarnberg, Mikael .....    | 288           |
| Garavaglia, Carla .....        | 100      | Müller, Mark .....         | 1                  | Thomson, Christopher ..... | 270           |
| Genero, Marcela .....          | 300, 357 | Nagappan, Nachiappan ..... | 120, 206           | Toda, Koji .....           | 360           |
| Godfrey, Sally .....           | 149      |                            |                    | Tosun, Ayse .....          | 318           |