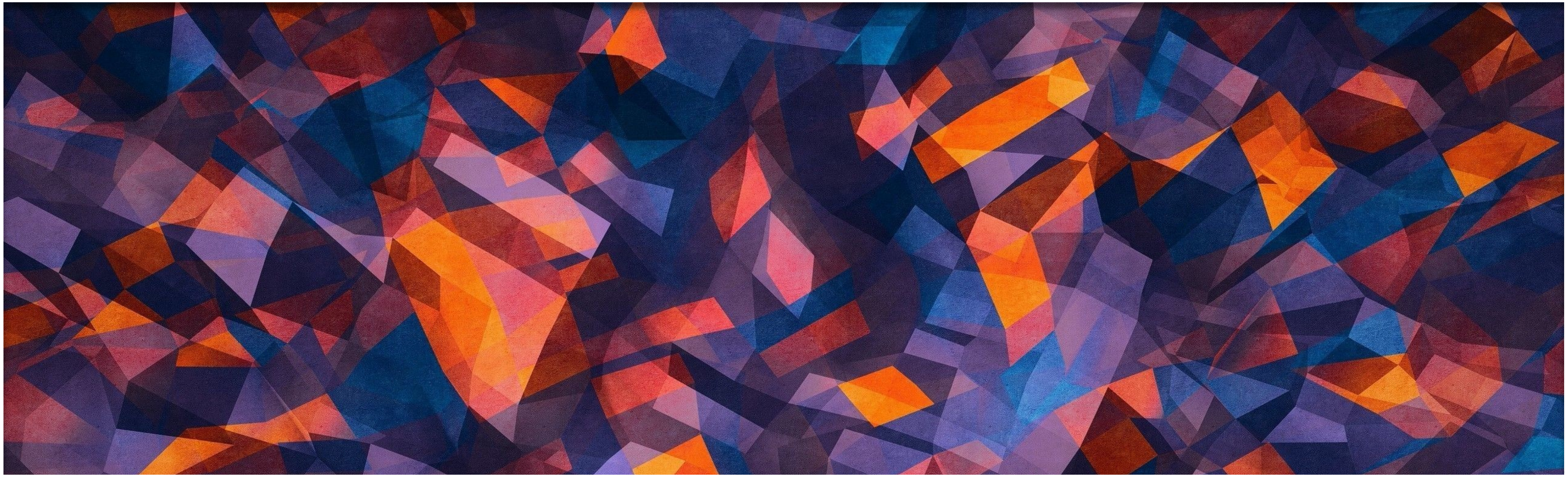
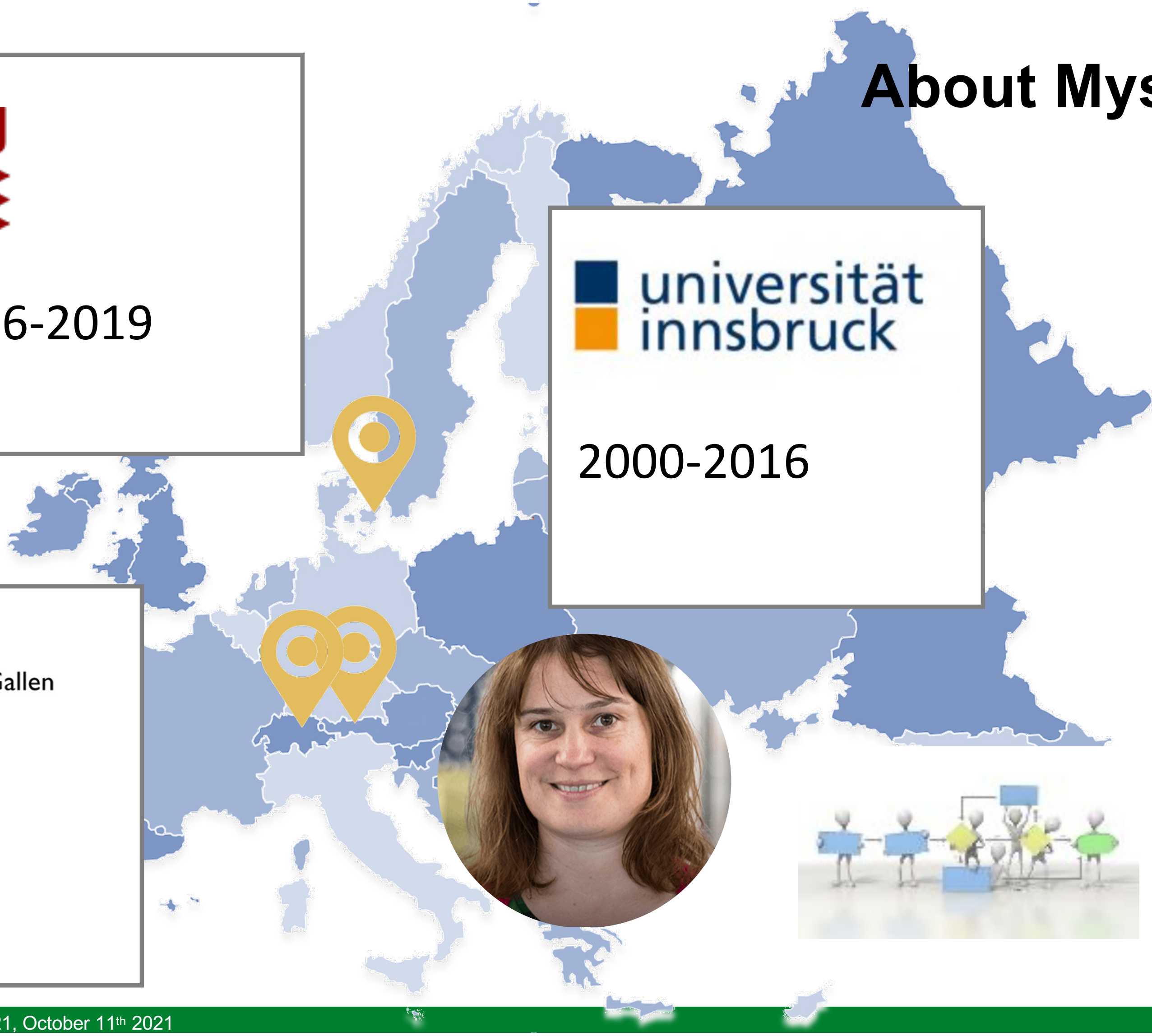
Universität St.Gallen



Evaluating the Outcome of Modeling Research

Prof. Dr. Barbara Weber

About Myself



DTU

2016-2019

universität innsbruck

2000-2016

University of St.Gallen

since 2019

Agenda

- Modeling Research as Design Science Research
- Types of Design Artifacts
- Three Cycle View of Design Science Research
- Methods for Evaluating Design Science Research
- Evaluation Framework for Design Science Research

Modeling Research

Modeling research deals with all aspects of modeling, from **languages** and **methods**, to **tools** and **applications**.

Modeling Research

Modeling research deals with all aspects of modeling, from **languages** and **methods**, to **tools** and **applications**.

Artifacts play a crucial role in modeling research.

Modeling Research as Design Science Research (DSR)

Modeling research deals with all aspects of modeling, from **languages** and **methods**, to **tools** and **applications**.

Artifacts play a crucial role in modeling research.

Design science research focuses on the **development** and **performance of (designed) artifacts** with the explicit intention of improving the functional performance of the artifact.

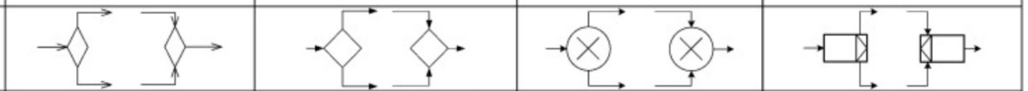
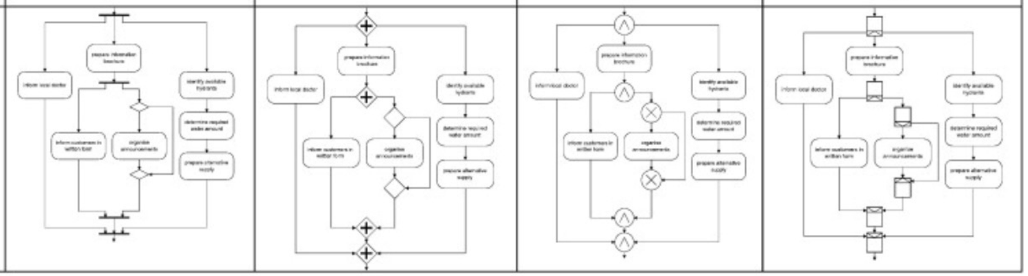
Types of Design Artifacts

Artifact Type	Description	Example
Construct	Vocabulary and symbols to define and understand problems and solutions; constitute the language to specify problems and solutions	Routing symbols (and other modeling primitives) in the process modeling domain (Figl et al. 2013)

Source: Hevner et al. 2004

Construct

Example: Routing Symbol Design

	R_{UML}	R_{BPMN}	R_{EPC}	R_{YAWL}
AND				
Outer Shape	narrow rectangle (bar)	symmetric diamond-shape	circle	rectangle
Inner Shape		internal marker (“+”)	logical marker for ‘and’ (“^”)	left- and right-sided open triangle
XOR				
Outer Shape	diamond-shape without internal marker	symmetric diamond-shape	circle	rectangle
Inner Shape		-	“X” marker	triangle
				

Routing symbols of four existing process modeling languages

Source: Figl et al. 2013

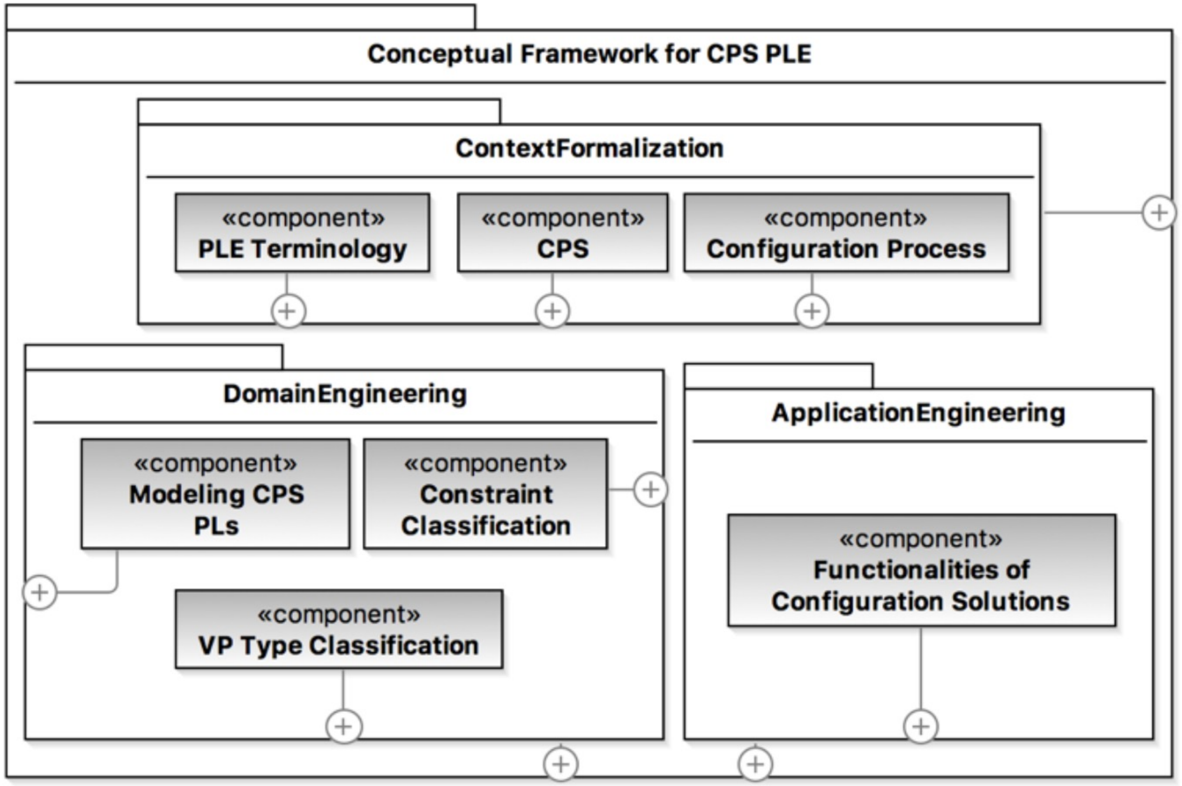
Types of Design Artifacts

Artifact Type	Description	Example
Construct	Vocabulary and symbols to define and understand problems and solutions; constitute the language to specify problems and solutions	Routing symbols (and other modeling primitives) in the process modeling domain (Figl et al. 2013)
Model	Designed representations of problem and/or solution;	Process models represented in BPMN; ER models; Conceptual framework to support automated product configuration in cyber physical systems (Safdar et al. 2020)

Source: Hevner et al. 2004

Model

Example: Framework for Product Configuration in Cyber Physical Systems (CPS)



Conceptual framework to support multi-stage and multi-step automated product configuration of CPSs. The framework includes classification of constraints and a list of automated functionalities of a CPS configuration solution.

Source: Safdar, S.A., Lu, H., Yue, T. et al. 2021

Types of Design Artifacts

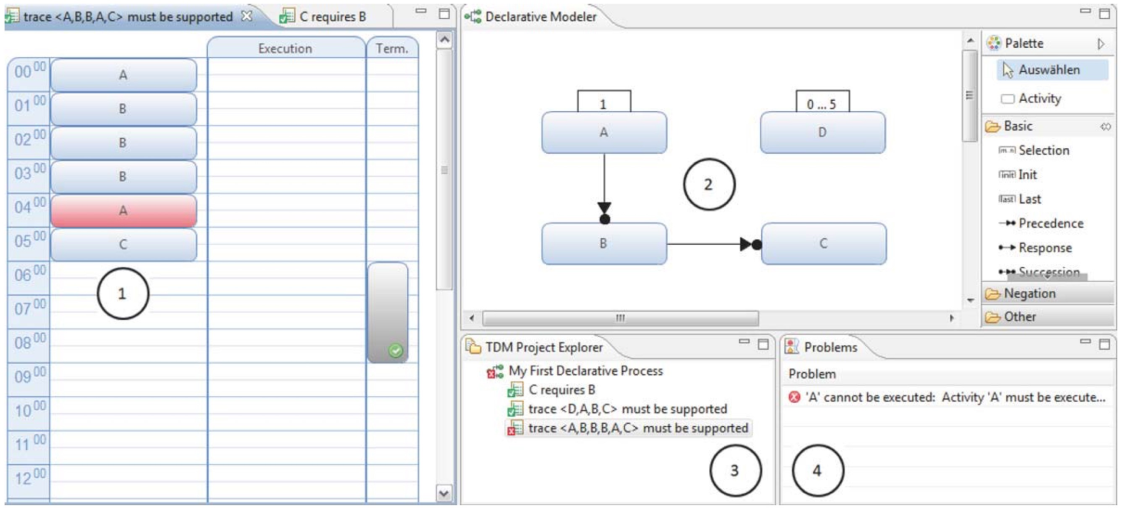
Artifact Type	Description	Example
Construct	Vocabulary and symbols to define and understand problems and solutions; constitute the language to specify problems and solutions	Routing symbols (and other modeling primitives) in the process modeling domain (Figl et al. 2013)
Model	Designed representations of problem and/or solution	Process models represented in BPMN; ER models; Conceptual framework to support automated product configuration in cyber physical systems (Safdar et al. 2020)
Method	Algorithms, practices, and recipes for performing a task	Test-driven modeling (Zugal et al. 2013); Structured Process Modeling Method (Claes et al. 2017)

Source: Hevner et al. 2004

Method

Instantiation

Example: Test-driven Modeling



To overcome problems in understanding and maintaining declarative process models **Test-driven Modeling** (a novel modeling method) has been proposed.

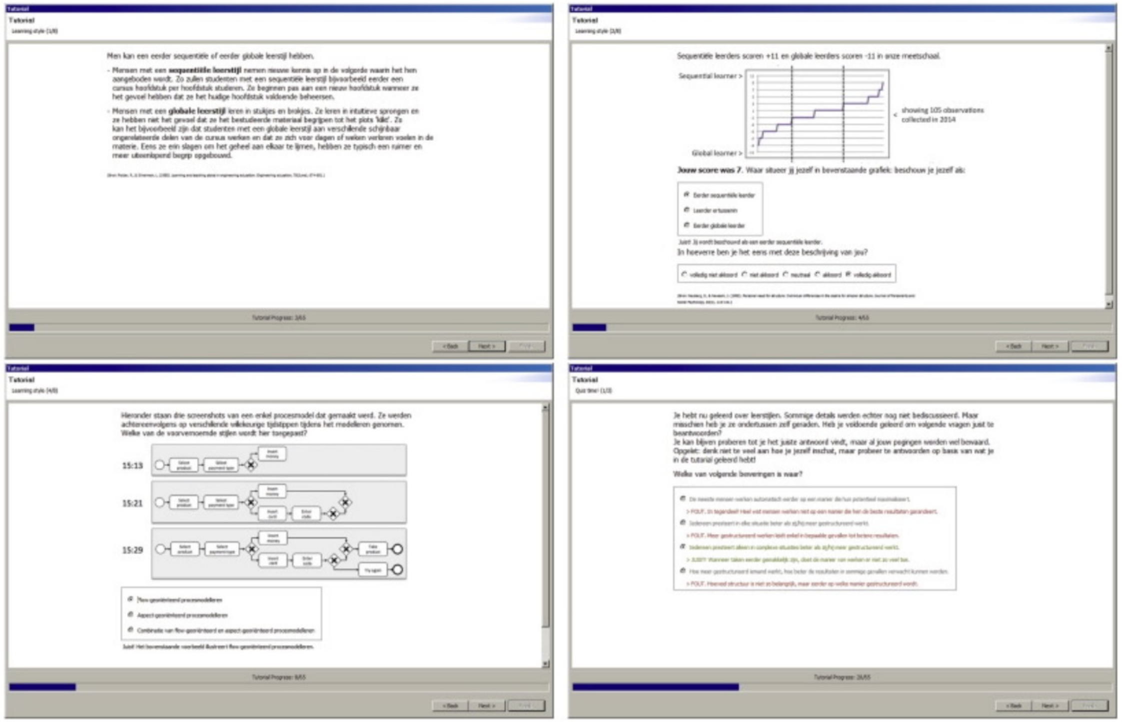
An implementation of the concepts of TDM are provided by **Test Driven Modeling Suite**.

Source: Zugal et al. 2013

Method

Instantiation

Example: Structured Process Modeling Method



Conceptual method to derive a modelers cognitive profile and the related optimal modeling strategy.

A **system** for operationally supporting the method through automated modeling strategy selection and training.

Source: Claes et al. 2017

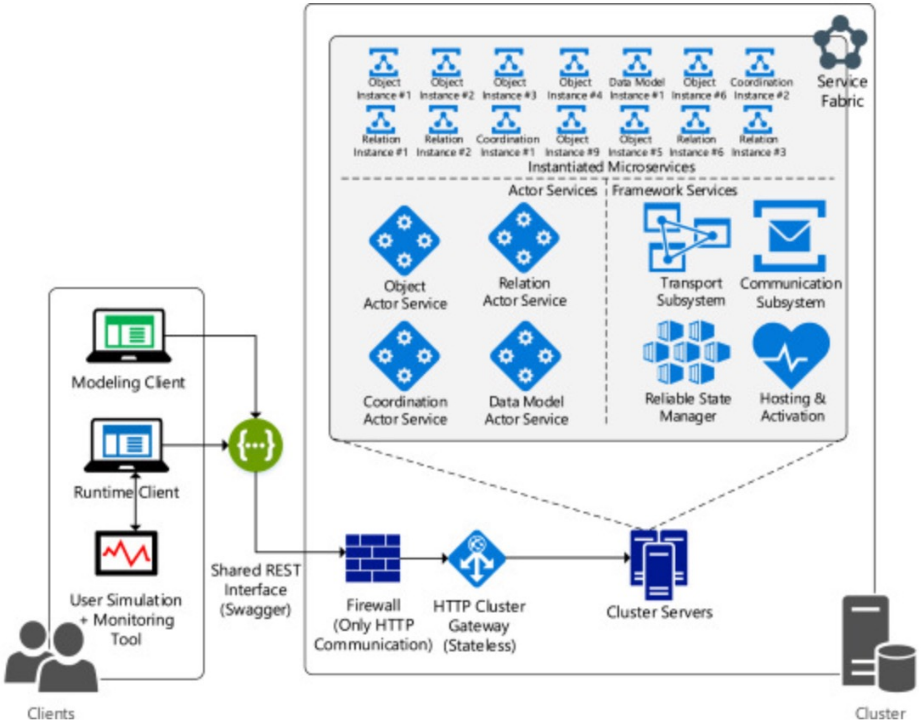
Types of Design Artifacts

Artifact Type	Description	Example
Construct	Vocabulary and symbols to define and understand problems and solutions; constitute the language to specify problems and solutions	Routing symbols (and other modeling primitives) in the process modeling domain (Figl et al. 2013);
Model	Designed representations of problem and/or solution	Process models represented in BPMN; ER models; Conceptual framework to support automated product configuration in cyber physical systems (Safdar et al. 2020)
Method	Algorithms, practices, and recipes for performing a task	Test-driven modeling (Zugal et al. 2013); Structured Process Modeling Method (Claes et al. 2017)
Instantiation	Physical realizations that act on the natural world (implemented systems, prototypes)	Prototype for runtime flexibility during data-centric and data-driven process execution (Andrews et al. 2019)

Method

Instantiation

Example: Run-time flexibility for data-aware and data-driven processes

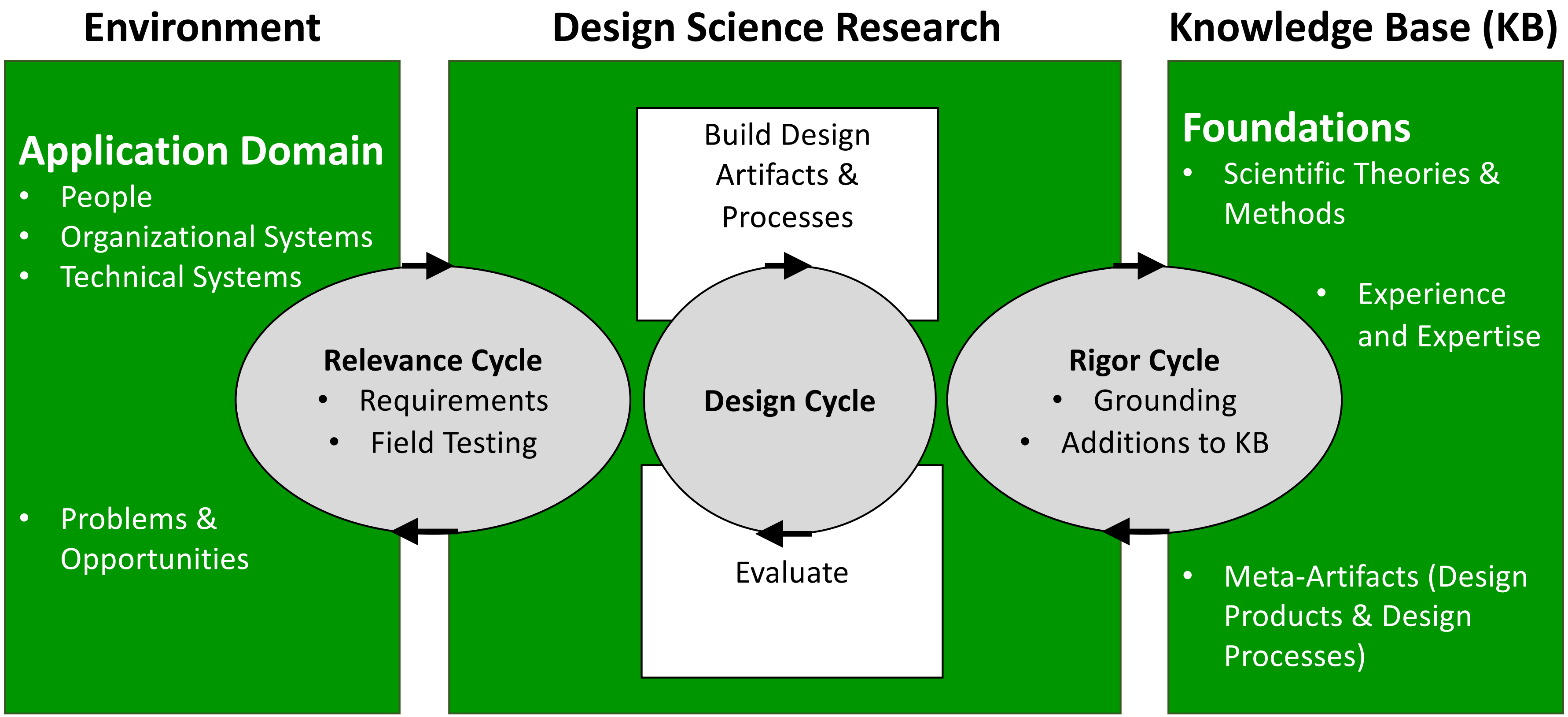


Concepts and algorithms for supporting ad-hoc changes during data-centric and data-driven process execution.

A **proof-of-concept** implementation and its application to various applications.

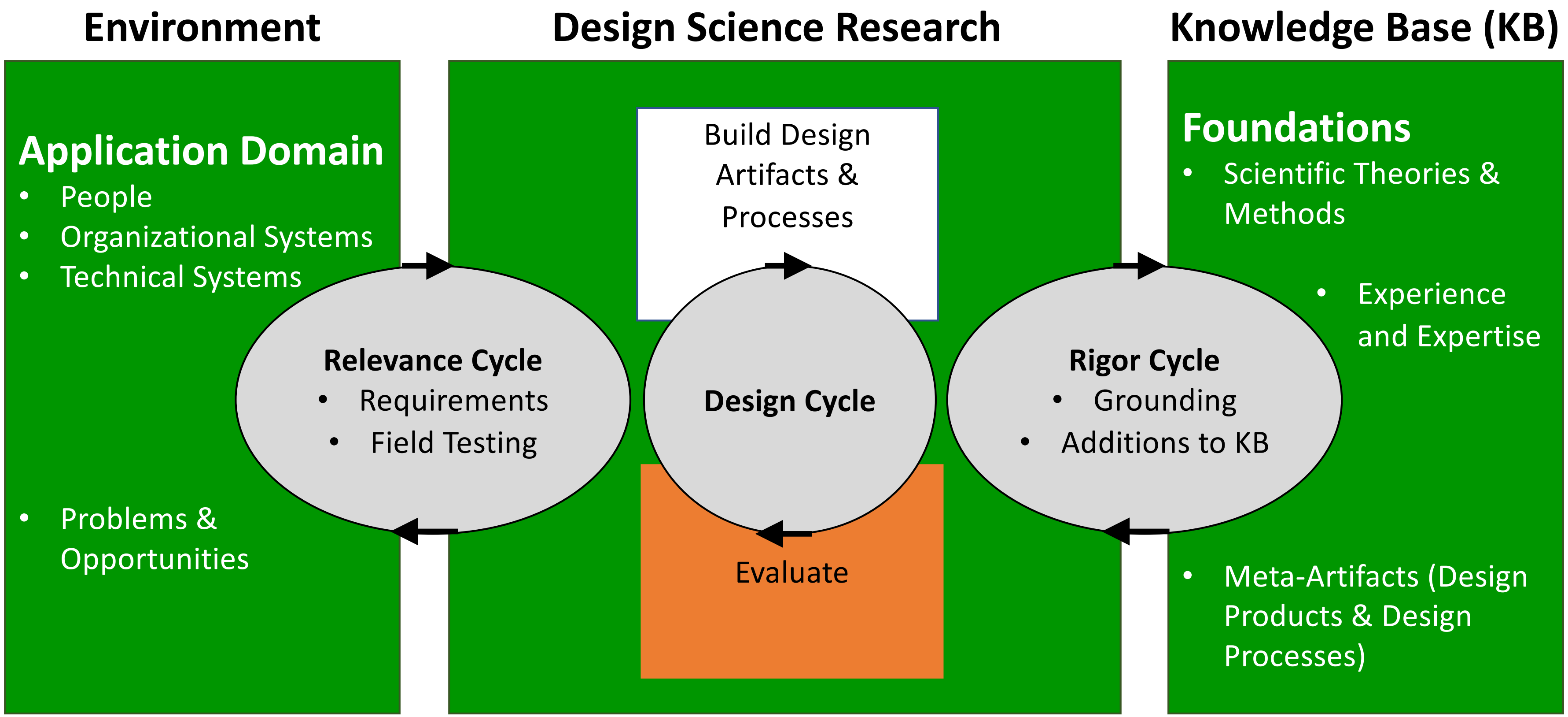
Source: Andrews et al. 2019

Design Science Research
Three Cycle View



Source: Hevner et al. 2007

Design Science Research
Three Cycle View



Source: Hevner et al. 2007

Methods for Evaluating Design Science Research (DSR)

Conceptual Evaluation

discusses the artifact's strength, weaknesses, and limitations

Empirical Evaluation

empirically evaluates the artifact using, for example, controlled experiments, case studies, action research, quantitative simulation, a benchmarking study

Source: https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/EngineeringResearch.md

Methods for Evaluating Design Science Research (DSR)

Conceptual Evaluation

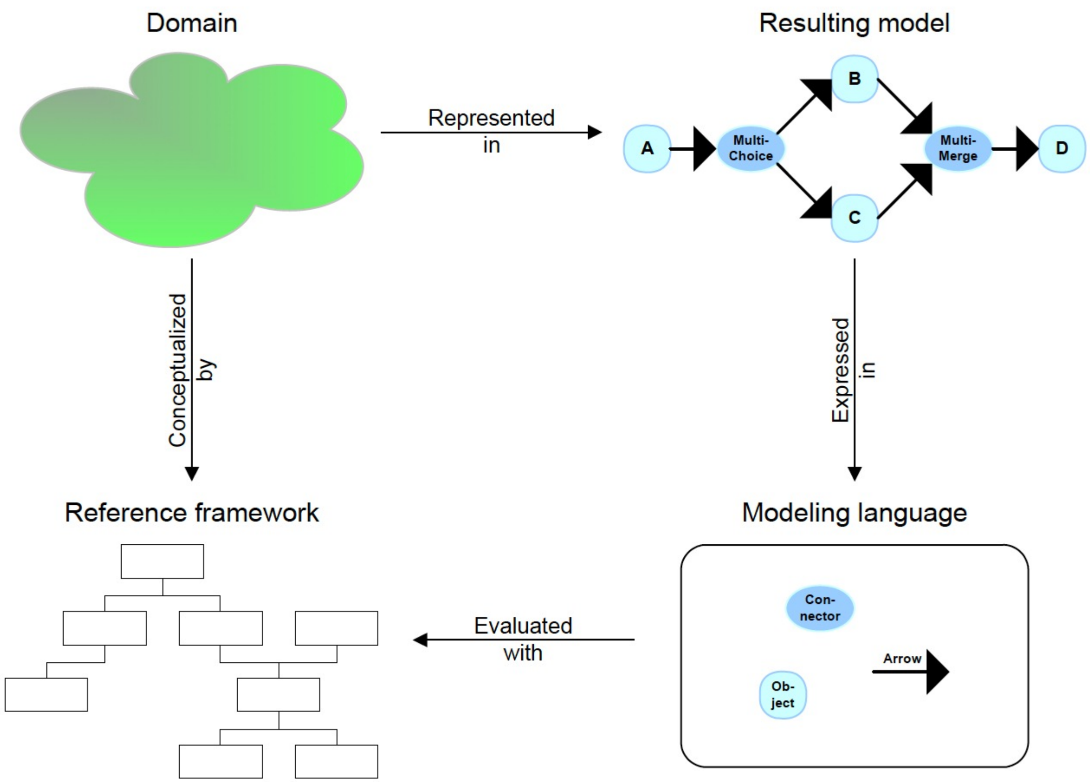
discusses the artifact's strength, weaknesses, and limitations

Empirical Evaluation

empirically evaluates the artifact using, for example, controlled experiments, case studies, action research, quantitative simulation, a benchmarking study

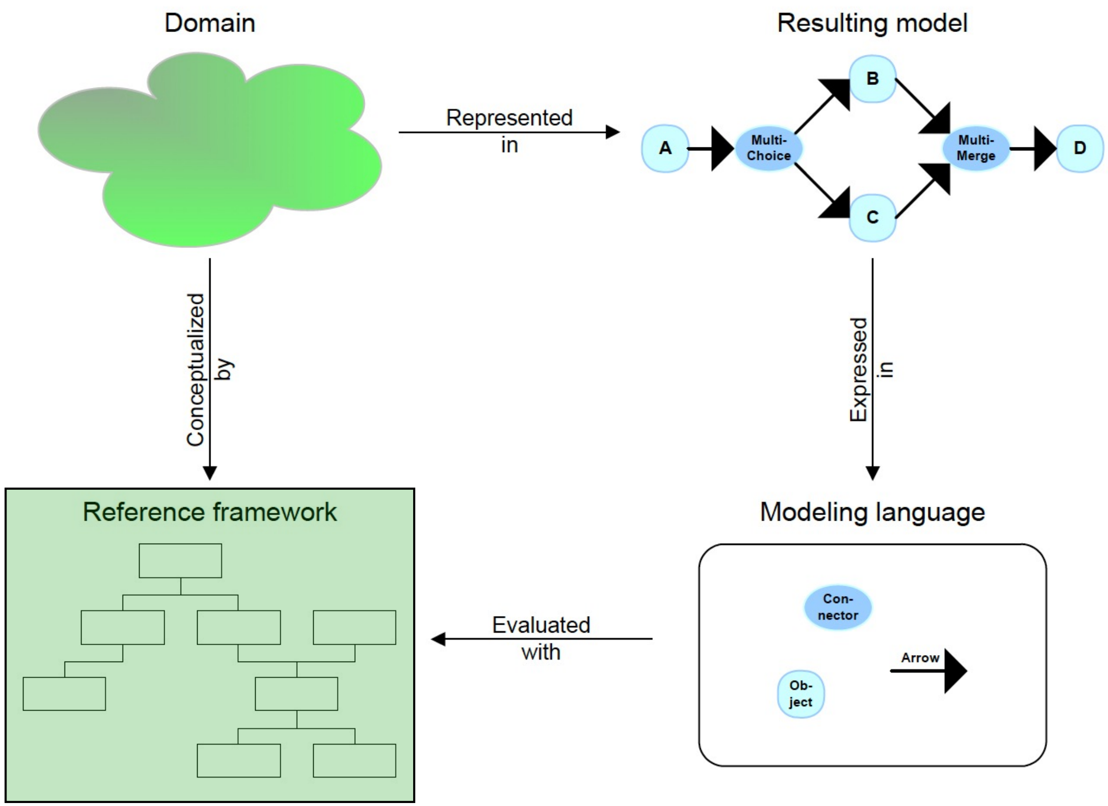
Source: https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/EngineeringResearch.md

Example: Conceptual Evaluation of Process Modeling Languages



Source: Recker et al. 2007

Example: Conceptual Evaluation of Process Modeling Languages



Source: Recker et al. 2007

Ontology

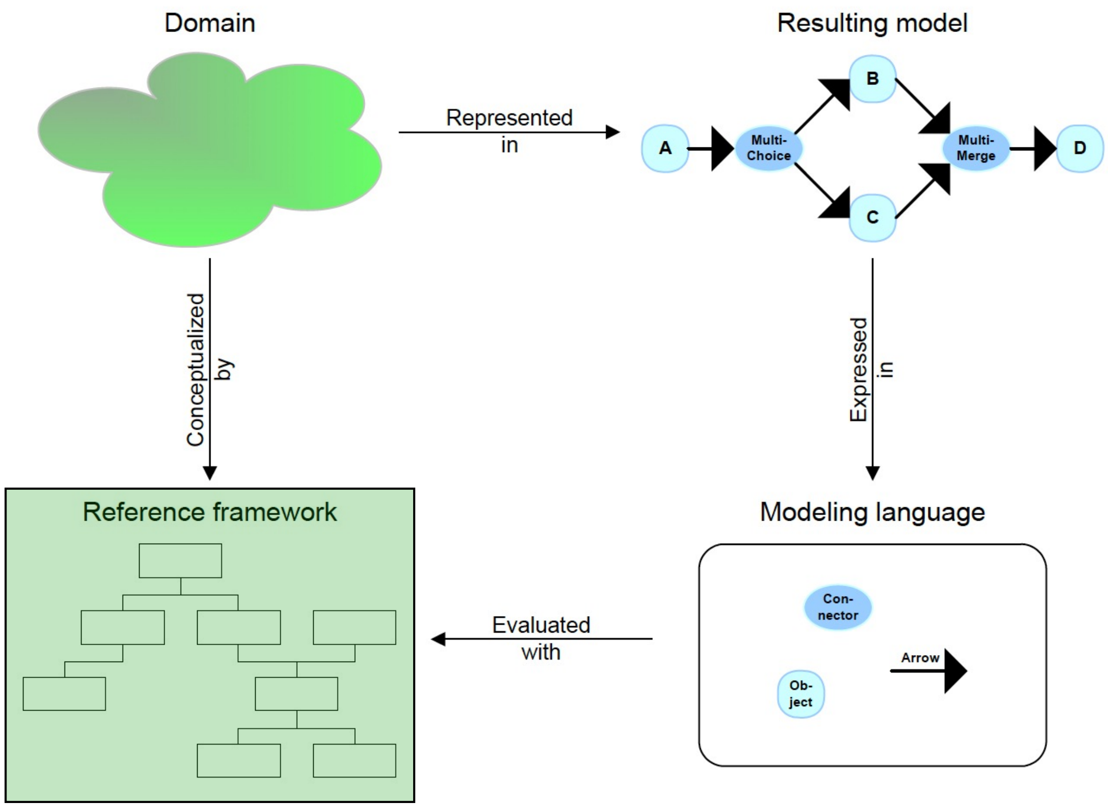
Bunge-Wand-Weber (BWW)
(Wand & Weber 1993)

Completeness of a description can be measured as:

- Construct deficit
- Construct overload
- Construct redundancy
- Construct excess

Source: Recker et al. 2007

Example: Conceptual Evaluation of Process Modeling Languages



Source: Recker et al. 2007

Patterns

Workflow patterns
(Russell et al. 2016)

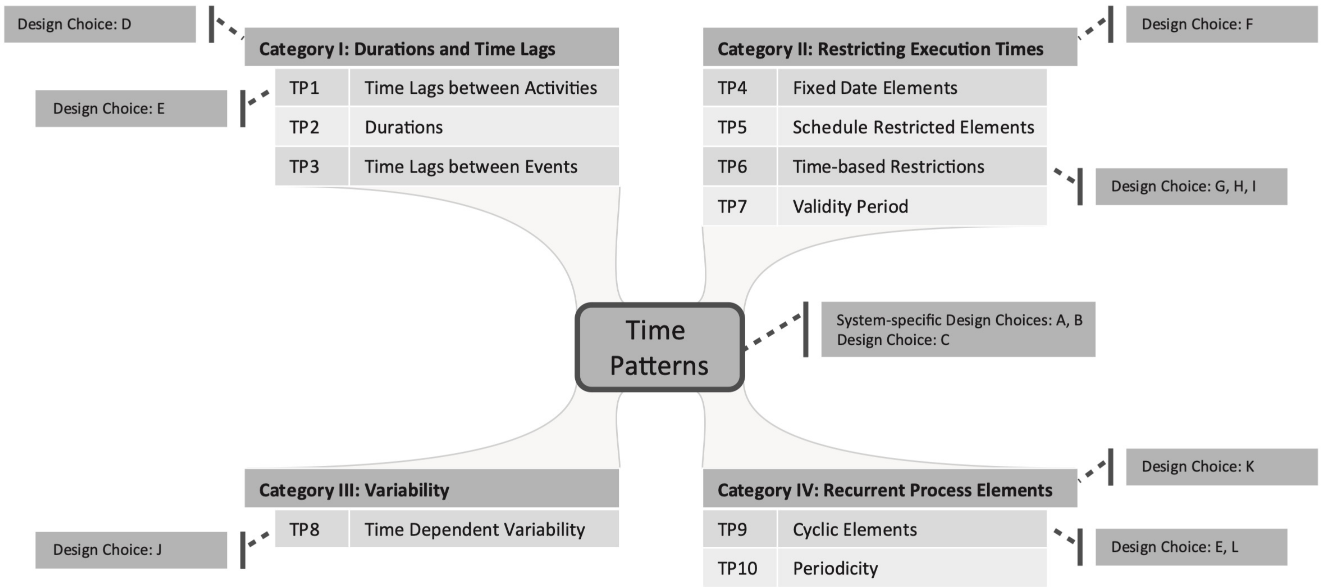
Change patterns
(Weber et al. 2008)

Time patterns
(Lanz et al. 2014)

VIVACE framework
(Ayora et al. 2015)

Example: Conceptual Evaluation of Modeling Languages

Overview of Identified Time-Patterns



Time patterns aim at the **comparison** of technologies for realizing time-and process-aware information systems. Moreover, they aim to provide a **reference for implementing** time support.

Source: Lanz et al. 2014

Example: Conceptual Evaluation of Modeling Languages

Overview of Identified Time-Patterns

Design Choice: D

Category I: Durations and Time Lags

TP1 Time Lags between Activities

Design Choice: E

Category II: Restricting Execution Times

TP4 Fixed Date Elements

Design Choice: F

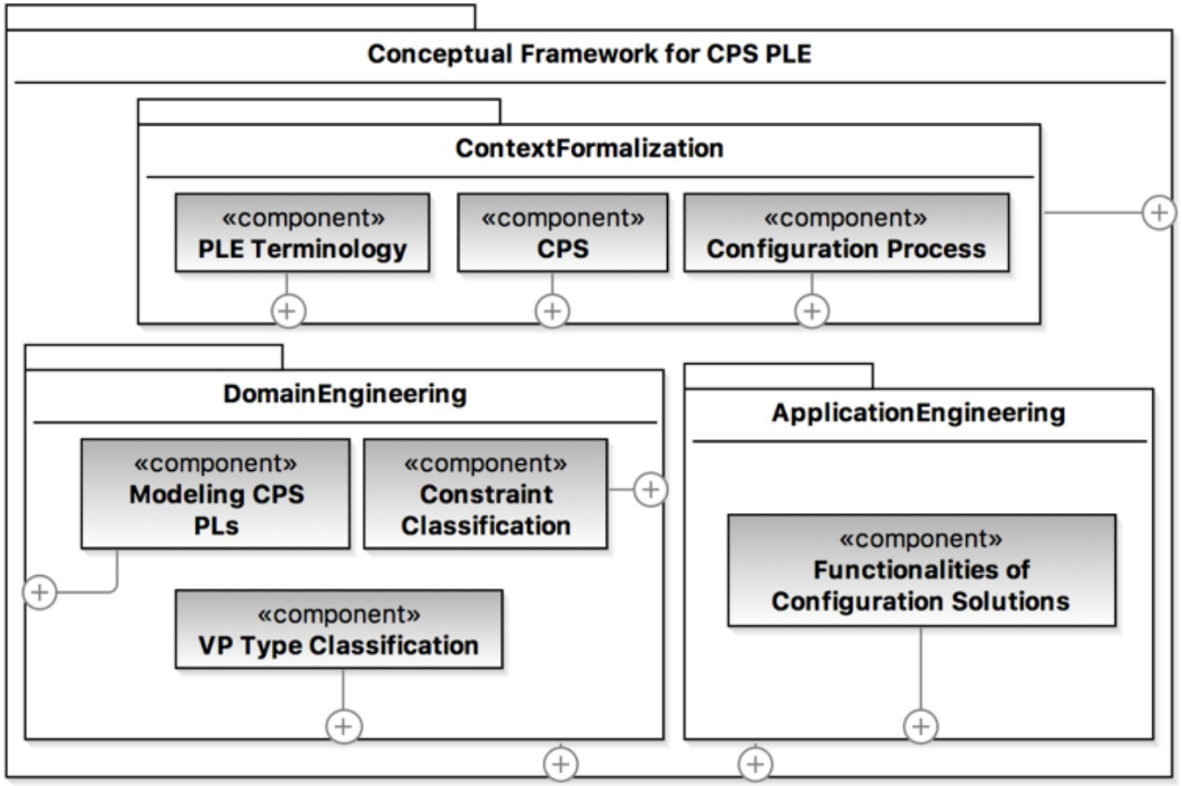
Patterns	Calendar Systems	Project Managememt	Academic			
	Microsoft Outlook 2010	Microsoft Project 2010	Bettini et al.	Combi et al.	Eder et al.	Zhuge et al.
System-specific Design Choices	A[b,c], B[a*,b*]	A[a,c], B[a,b,c]	A[a,b?,c?], B[a,b,c]	A[a], B[a,b]	A[a,b,c], B[a*]	A[a,c]
Category I: Durations and Time Lags						
TP1 - Time Lags between Activities	—	D[a,b], E[a,b,c,d]	D[a,b,c], E[a,b,c,d]	D[a,b,c], E[a,b,c,d]	D[a,b,c], E[d]	D[a,b,c], E[c*]
TP2 - Durations	C[a], D[b]	C[a,c], D[b]	C[a,c], D[a,b,c]	C[a,c], D[a,b,c]	C[a,c], D[b]	C[a], D[a,b,c]
				D[a*,b*,c*]	—	—
				C[a], F[a,b*,c,d]	C[a], F[c]	C[a], F[b]
				C[a], F[a,b]	C[a], F[c]	—
				—	—	—
				C[a], F[a,d]	C[a?], F[c?,d?]	C[a?], F[b?]
				—	—	—
				[a,b,c], E[a?,b?,c?,d?], K[a], L[b]	—	—
				M[a,b,c]	—	—

Patterns	Standards		Commercial				
	BPMN 2.0	WS-BPEL4People 2.0	IBM Websphere Integration Developer 6.1	WebSphere Lombardi Edition 7.1	AristaFlow 1.0.1	Intalio 6.0.3	TIBCO Business Studio 3.4.2
System-specific Design Choices	A[a,b?,c], B[a?,b?,c?]	A[a,c], B[a?,b?,c?]	A[a,c], B[a]	A[a,c], B[a,c*]	A[a,c], B[a,b]	A[a,c], B[a]	A[a,c], B[a]
Category I: Durations and Time Lags							
TP1 - Time Lags between Activities	D[a,b*,c*], E[c*]	D[a,b*,c*], E[c]	D[a,b,c], E[c,d*]	D[a,b*,c*], E[a*,b*,c,d*]	D[b], E[c*,d*]	D[a,b*,c*], E[c]	D[a,b*,c*], E[c*]
TP2 - Durations	C[a,c*], D[b]	C[a,c*], D[b]	C[a,c], D[b]	C[a,c], D[b]	C[a], D[b]	C[a*,c*], D[b]	C[a,c*], D[b]
TP3 - Time Lags between Events	D[a,b*,c*]	D[a]	D[a]	D[a*]	—	D[a]	D[a,b*,c*]
Category II: Restrictions of Process Execution Points							
TP4 - Fixed Date Elements	C[a,b*], F[a,b?,d]	C[a], F[a,d]	C[a], F[a,b*,d*]	C[a,c*], F[a*,d]	C[a], F[b*,d]	—	C[a,c*], F[a,b*,d]
TP5 - Schedule Restricted Elements	—	—	—	C[a*,c*], F[a,b]	—	—	—
TP6 - Time Based Restrictions	—	—	—	—	—	—	—
TP7 - Validity Period	—	—	C[c], F[a]	—	—	—	—
Category III: Variability							
TP8 - Time Dependent Variability	J[a,b*]	J[a,b*]	J[a,b*]	J[a]	J[a]	J[a,b*]	J[a,b*]
Category IV: Reoccurring Process Elements							
TP9 - Cyclic Elements	D[a*], E[c*], K[a], L[a,b]	D[a*], E[c], K[a], L[a,b]	D[a*], E[a*,c], K[a], L[a,b]	D[a*], E[c*], K[a], L[a,b]	D[b], E[c*,d*], K[a], L[a,b]	D[a*], E[c], K[a], L[a,b]	D[a*], E[c*], K[a], L[a]
TP10 - Periodicity	—	—	—	—	—	—	—

Time patterns aim at the **comparison** of technologies for realizing time-and process-aware information systems. Moreover, they aim to provide a **reference for implementing** time support.

Source: Lanz et al. 2014

Example: Framework for Product Configuration in Cyber Physical Systems (CPS)



Conceptual framework to support multi-stage and multi-step automated product configuration of CPSs

Framework can serve as guide to researchers and practitioners to **evaluate an existing CPS-specific PLE solution** or **devise a new one.**

Source: Safdar, S.A., Lu, H., Yue, T. et al. 2021

Methods for Evaluating Design Science Research (DSR)

Conceptual Evaluation

discusses the artifact's strength, weaknesses, and limitations

Empirical Evaluation

empirically evaluates the artifact using, for example, controlled experiments, case studies, action research, quantitative simulation, a benchmarking study

Source: https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/EngineeringResearch.md

Empirical Methods for Evaluating DSR

Controlled Experiment

human participants use the artifact

Case Study

researchers observe a real organization using the artifacts

Action Research

researchers intervene in a real organization using the artifact

Quantitative Simulation

artifact is assessed (usually against a competing artifact) in an artificial environment

Benchmarking Study

artifact is assessed using one or more benchmarks

Source: https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/EngineeringResearch.md

Variables for the Evaluation of DSR Artifacts

Variable	Value				
Approach	Qualitative		Quantitative		
Artifact Focus	Technical	Organizational		Strategic	
Artifact Type	Construct	Model	Method	Instantiation	Theory
Epistemology	Positivism		Interpretivism		
Function	Knowledge function	Control function	Development function	Legitimization function	
Method	Action research	Case study	Field experiment	Formal proofs	
	Controlled experiment	Prototype	Survey		
Object	Artifact		Artifact construction		
Ontology	Realism		Nominalism		
Perspective	Economic	Deployment	Engineering	Epistemological	
Position	Externally		Internally		
Reference Point	Artifact against research gap	Artifact against real world	Research gap against real world		
Time	Ex ante		Ex post		

Source: Cleven et al. 2009

Evaluating Design Science Research
Two Distinct Goals

Goal Type I

Demonstrate that a very new artifact works, i.e., a solution to an unsolved problem was found.

Goal Type II

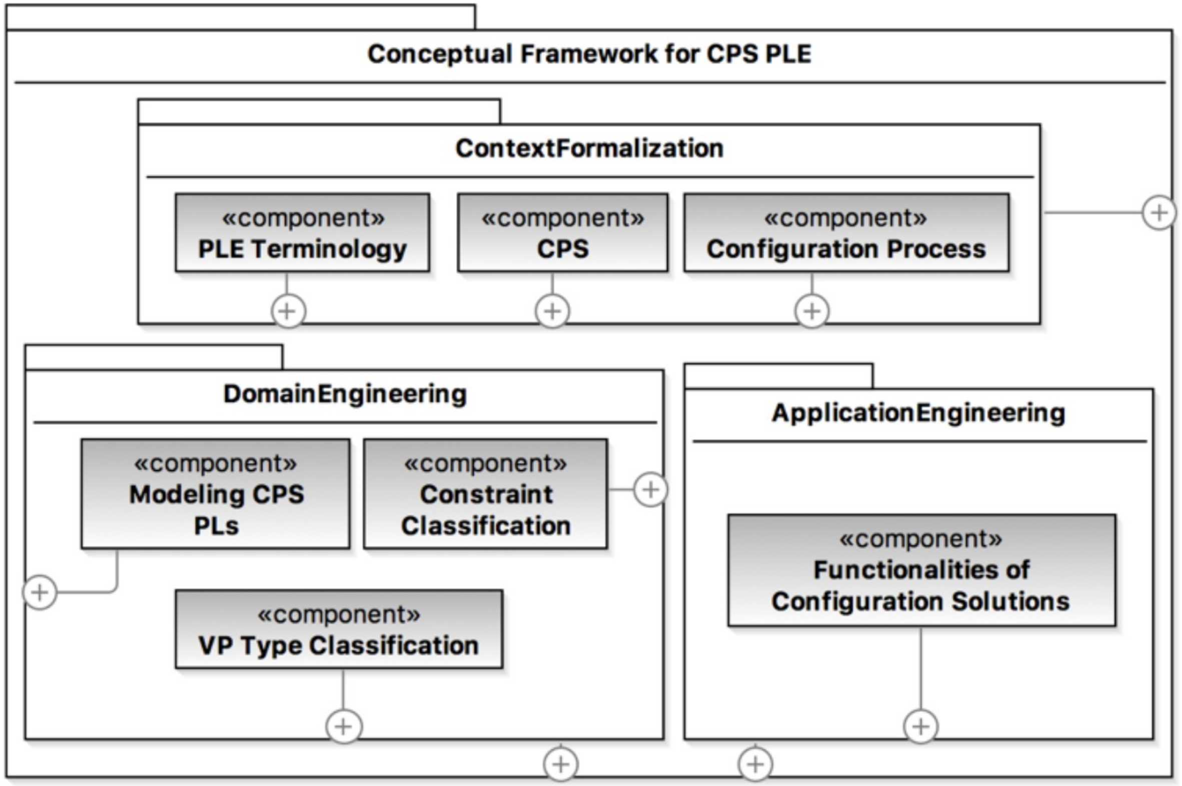
Demonstrate that the design artifact works better than existing solutions, e.g., the artifact solves a problem much more efficiently and with fewer resources.

Source: Mettler et al. 2014

Goal Type I

Case Study

Example: Framework for Product Configuration in Cyber Physical Systems (CPS)



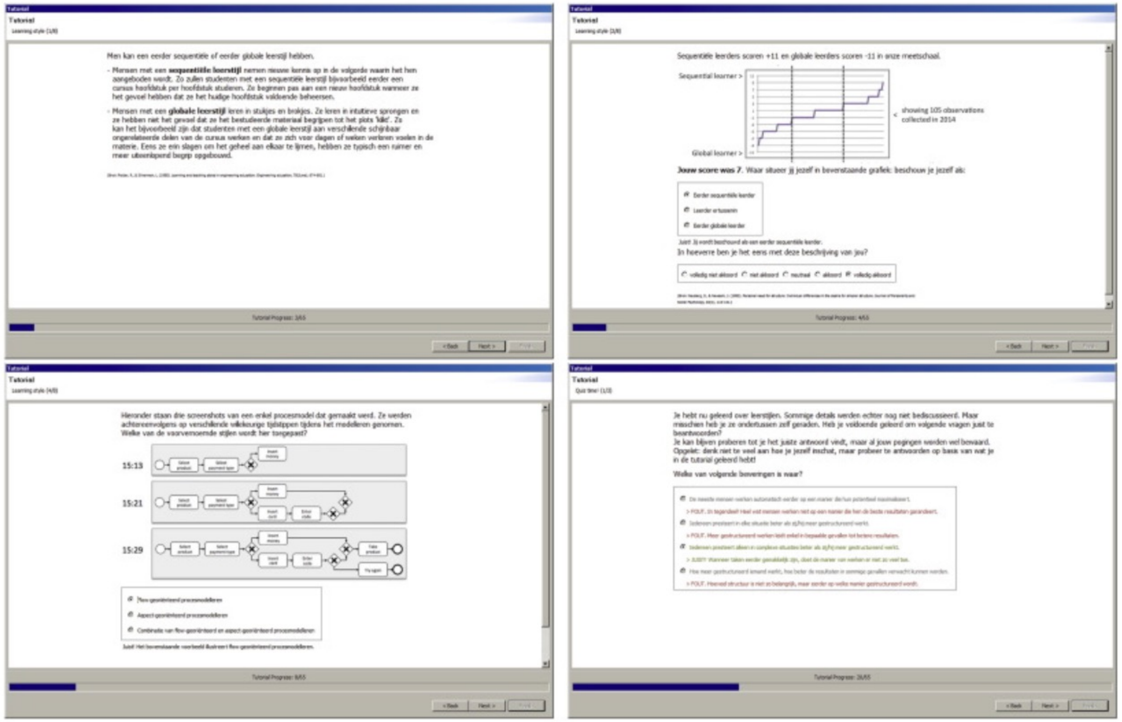
Assess if the framework provides the support for **capturing** and **managing** **commonalities**, **variabilities**, and **constraints** in the domain engineering phase as well as the **support for automation of configuration** in the application engineering phase.

Source: Safdar, S.A., Lu, H., Yue, T. et al. 2021

Goal Type I

Controlled Experiment

Example: Structured Process Modeling Method



Assess if **selecting** and **training** modelers in their **optimal process modeling strategy** (i.e., the proposed Structured Modeling Method) improves modeling performance.

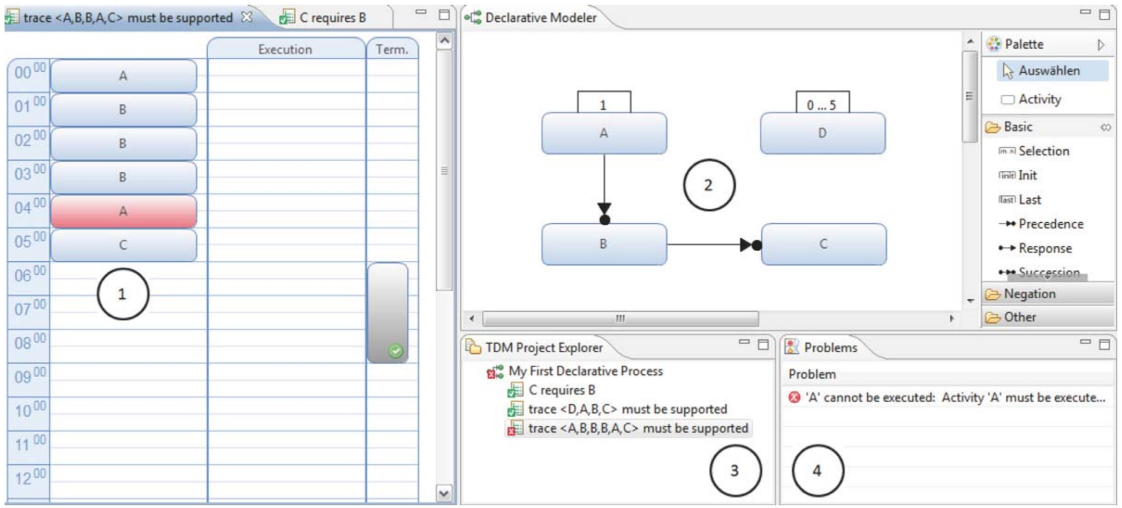
Source: Claes et al. 2017

Goal Type I

Case Study

Controlled Experiment

Example: Test-driven Modeling



Source: Zugal et al. 2013

Assess the impact of Test-driven Modeling on **communication behavior** (communication between domain experts and model builders) using a case study.

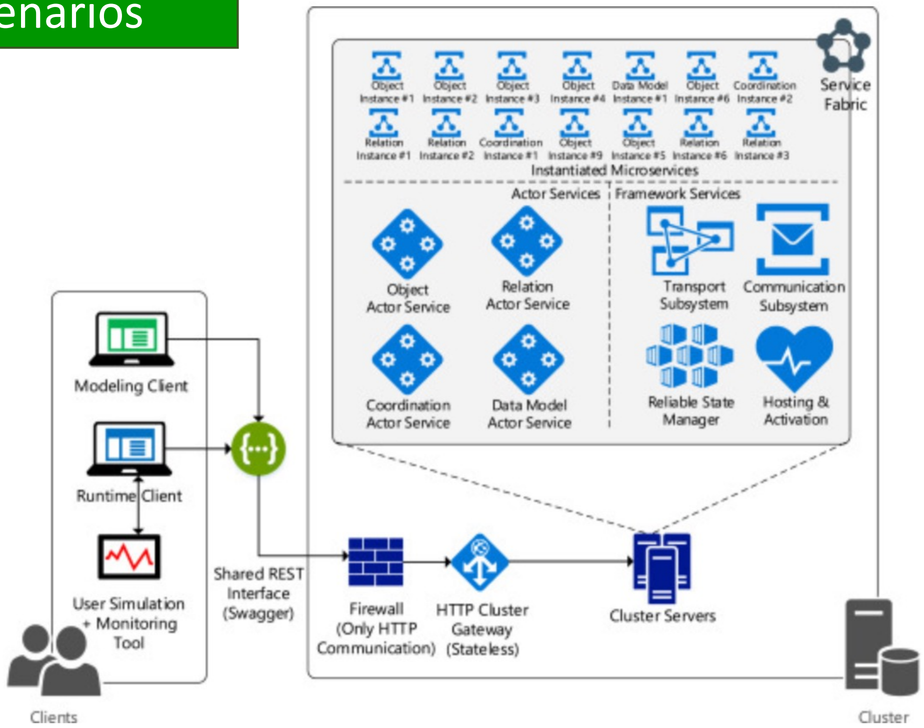
Assess the impact of Test-driven Modeling on **maintainability** using a controlled experiment.

Goal Type I

Prototype

Scenarios

Example: Run-time flexibility for data-aware and data-driven processes



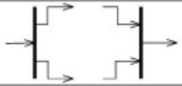
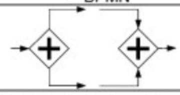
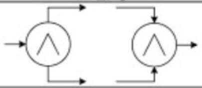
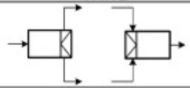
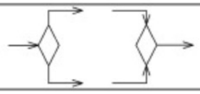
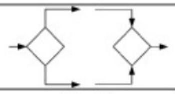
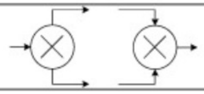
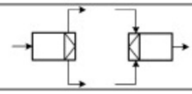
Assess the **feasibility** of ad-hoc changes without disruptions.

Show that changes are **possible** concerning every aspect of a process model.

Evaluate the **scalability** of the approach.

Source: Andrews et al. 2019

Example: Routing Symbol Design

	R_{UML}	R_{BPMN}	R_{EPC}	R_{YAWL}
AND				
Outer Shape	*narrow rectangle (bar)*	*symmetric diamond-shape*	*circle*	*rectangle*
Inner Shape		*internal marker (“+”)*	*logical marker for ‘and’ (“^”)*	*left- and right-sided open triangle*
XOR				
Outer Shape	*diamond-shape without internal marker*	*symmetric diamond-shape*	*circle*	*rectangle*
Inner Shape		*-*	*“X” marker*	*triangle*
				

Comparing the effect of different **routing symbol designs** of four different pre-existing modeling languages in the context of process model comprehension tasks.

Source: Figl et al. 2013

Empirical Methods for Evaluating DSR

Controlled Experiment

human participants use the artifact

Case Study

researchers observe a real organization using the artifacts

Action Research

researchers intervene in a real organization using the artifact

Quantitative Simulation

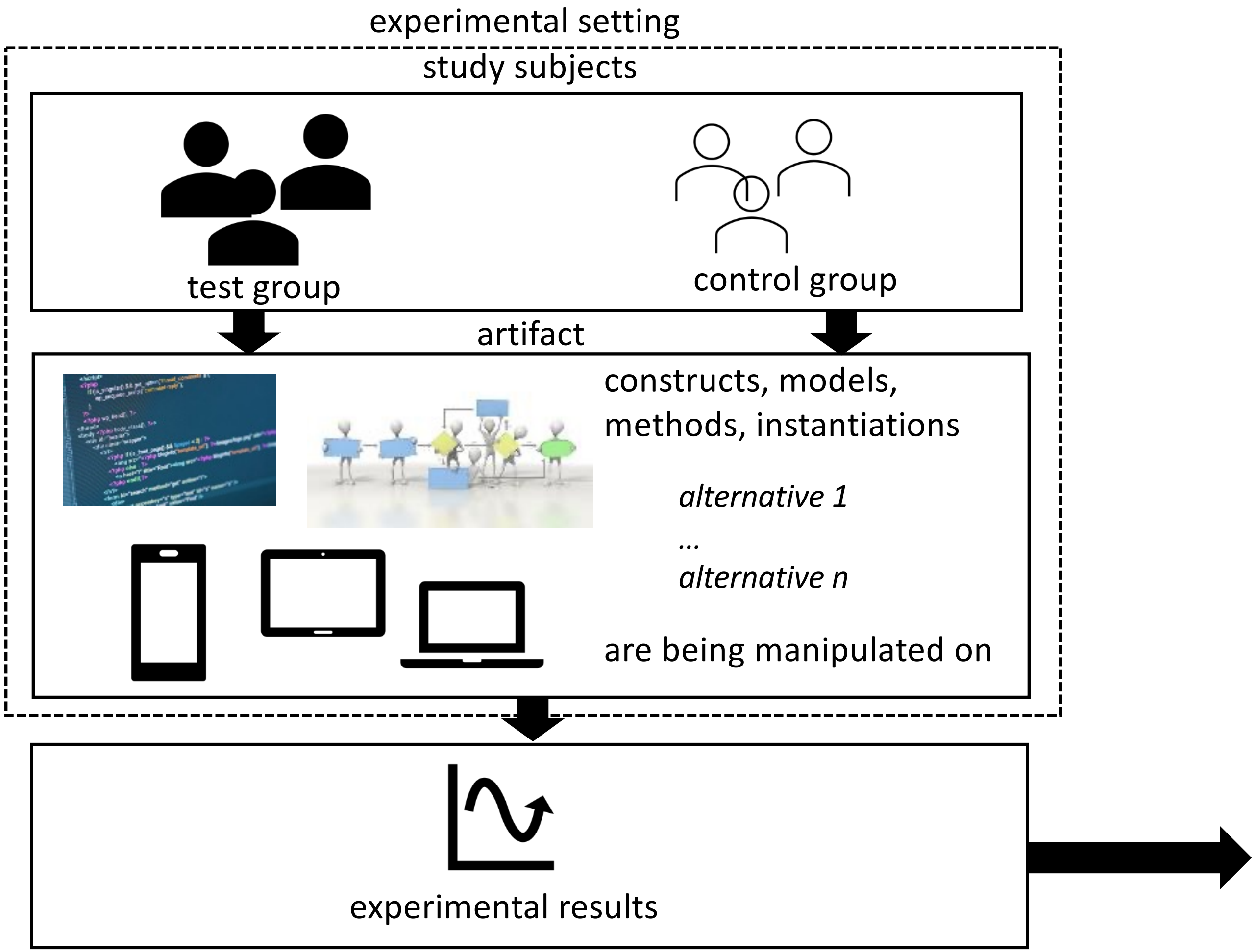
artifact is assessed (usually against a competing artifact) in an artificial environment

Benchmarking Study

artifact is assessed using one or more benchmarks

Source: https://github.com/acmsigsoft/EmpiricalStandards/blob/master/docs/EngineeringResearch.md

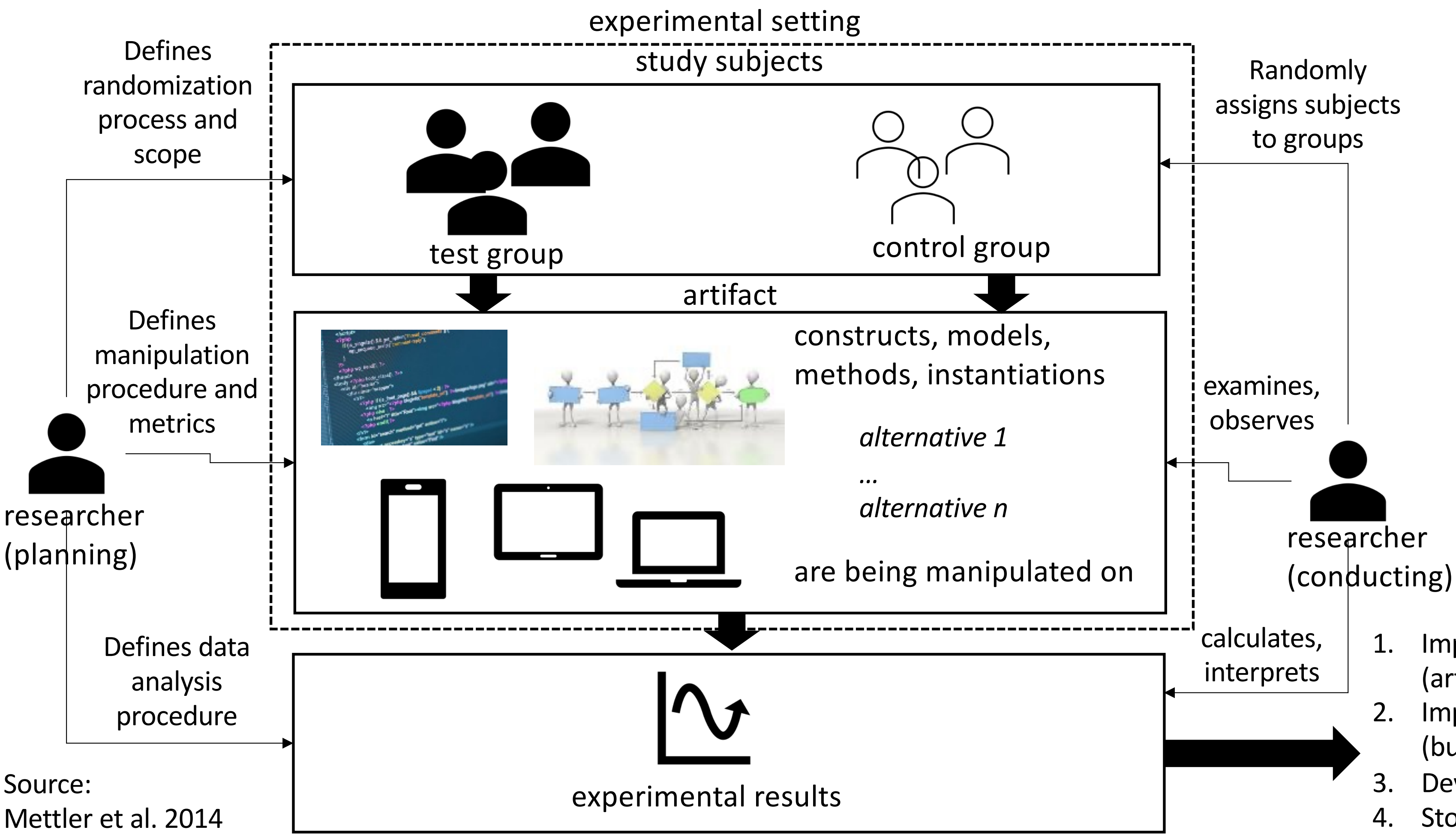
Use of Experiments in Design Science Research



1. Improve design product (artifact)
2. Improve design process (build/test cycle)
3. Develop design theory
4. Stop (build/test cycle)

Source:
Mettler et al. 2014

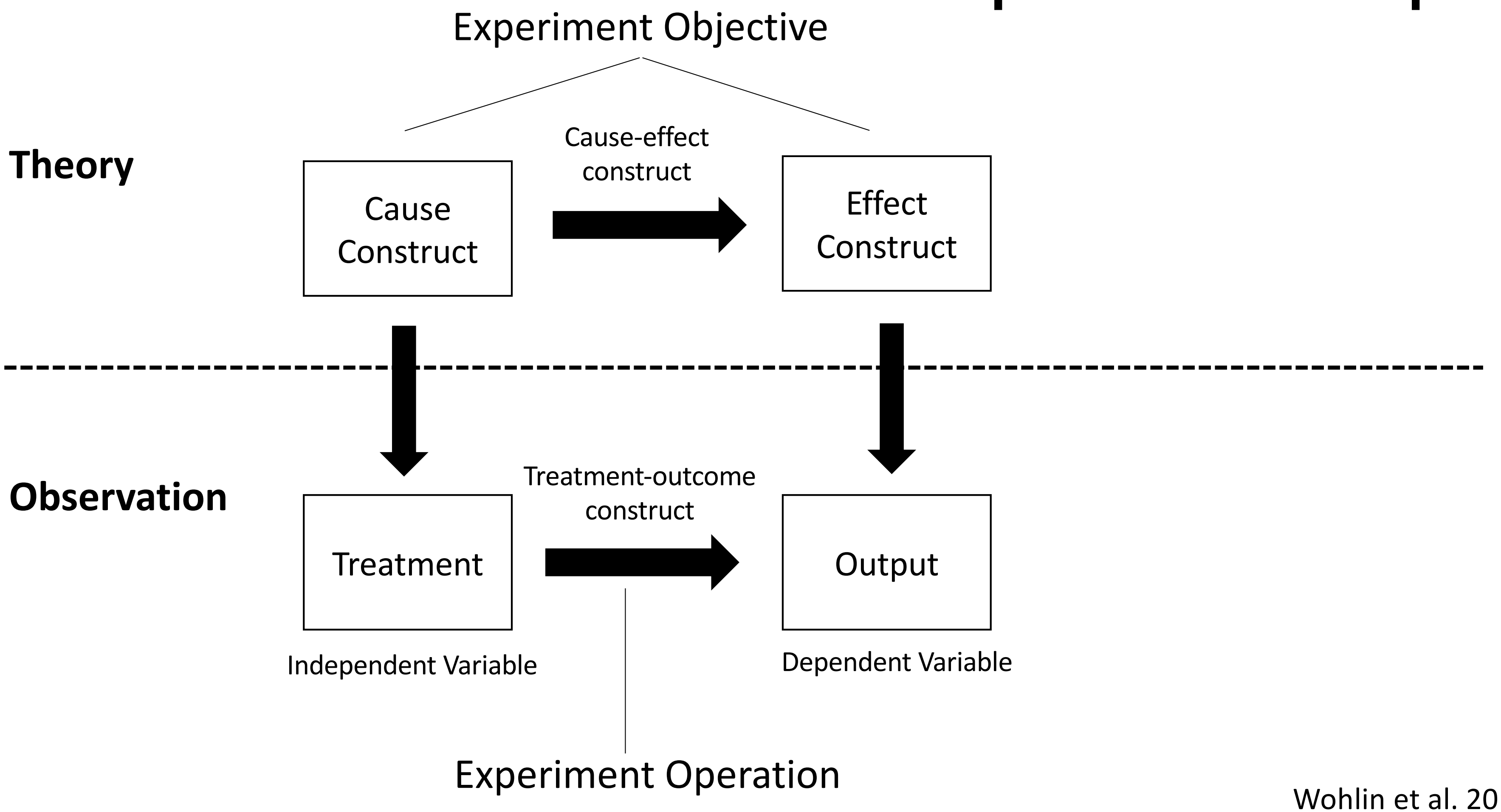
Use of Experiments in Design Science Research



1. Improve design product (artifact)
2. Improve design process (build/test cycle)
3. Develop design theory
4. Stop (build/test cycle)

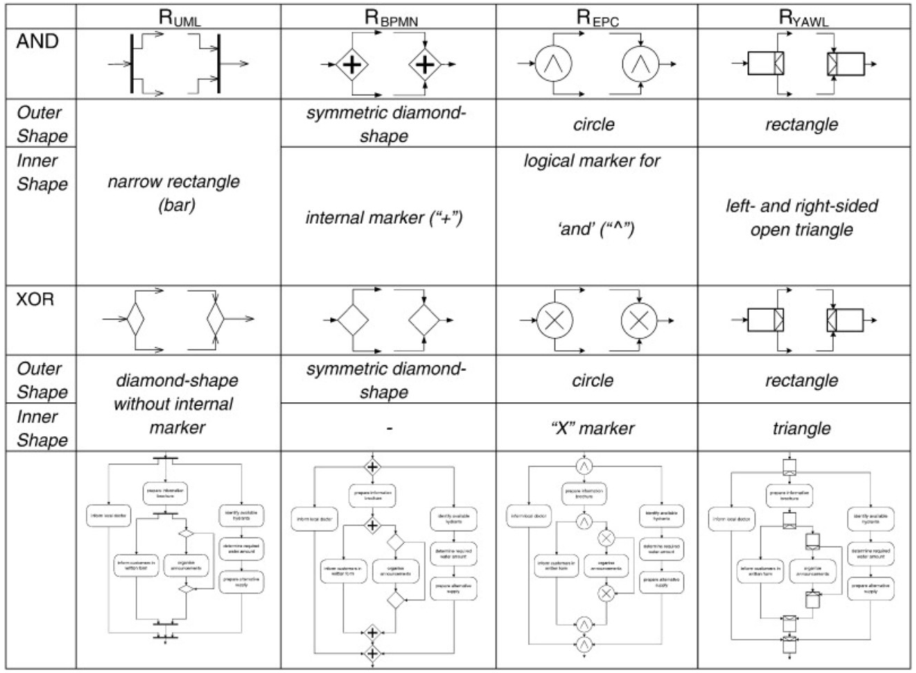
Source:
Mettler et al. 2014

Experiment Principles



Wohlin et al. 2012

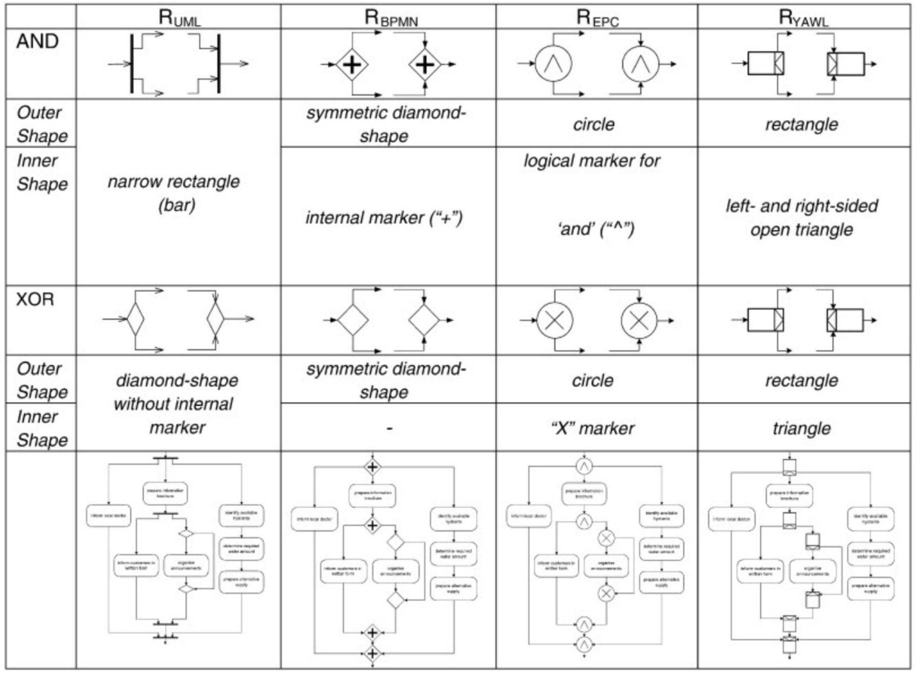
Example: Routing Symbol Design



Effect of **routing symbol design** (i.e., perceptual discriminability, pop out, semantic transparency, and aesthetic) design during model comprehension task on model comprehension **accuracy**, **efficiency**, and **perceived cognitive load**.

Source: Figl et al. 2013

Example: Routing Symbol Design



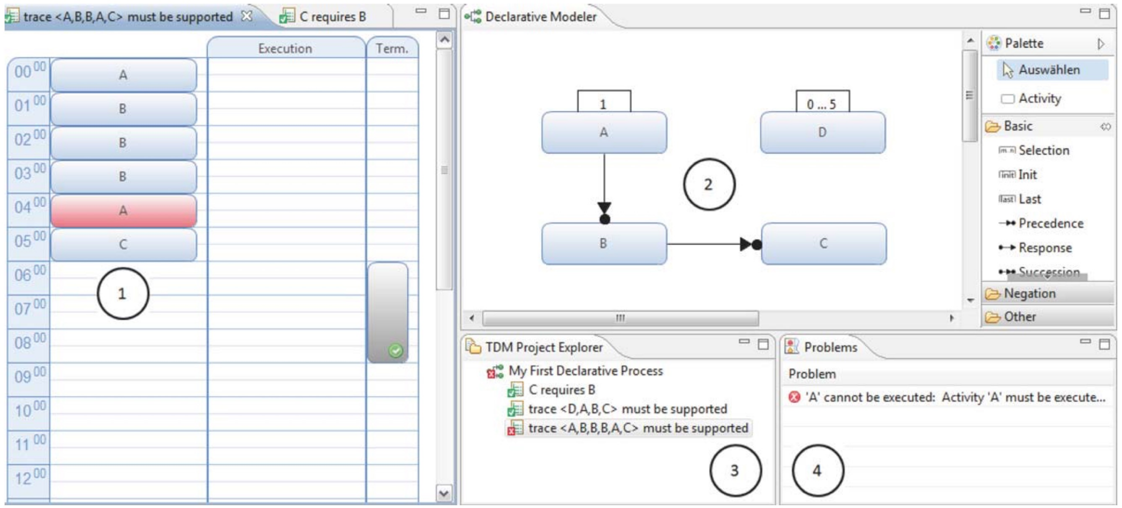
Effect of **routing symbol design** (i.e., perceptual discriminability, pop out, semantic transparency, and aesthetic) design during model comprehension task on model comprehension **accuracy**, **efficiency**, and **perceived cognitive load**.

Results: Design principles related to

- Perceptual discriminability and pop out improve comprehension accuracy
- Semantic transparency and aesthetic design of symbols lower perceived difficulty of comprehension

Source: Figl et al. 2013

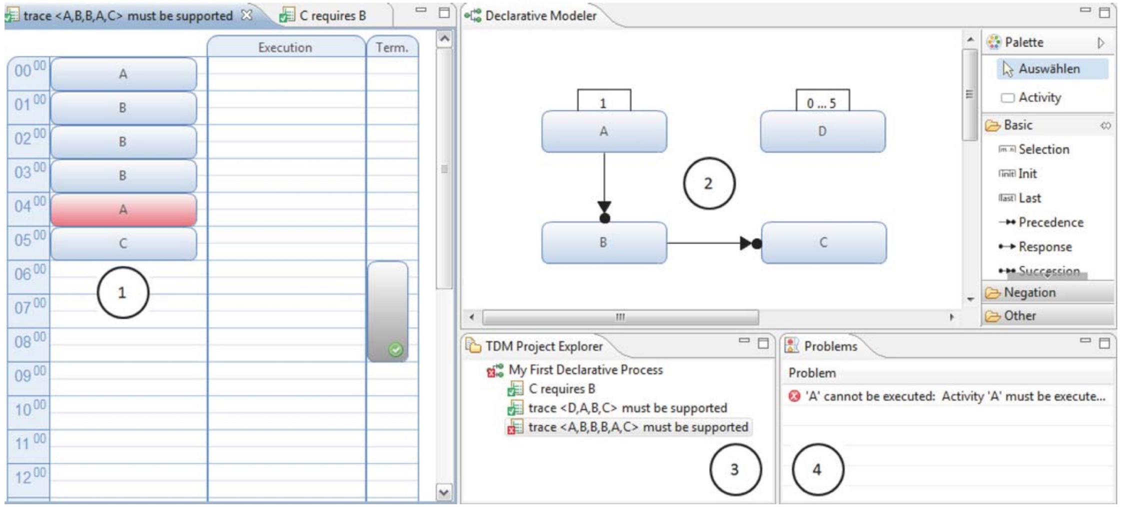
Example: Test-driven Modeling



Effect of **Test-driven modeling** on **perceived cognitive load, perceived quality**, and **quality of the adapted models** in the context of maintainability tasks.

Source: Zugal et al. 2013

Example: Test-driven Modeling



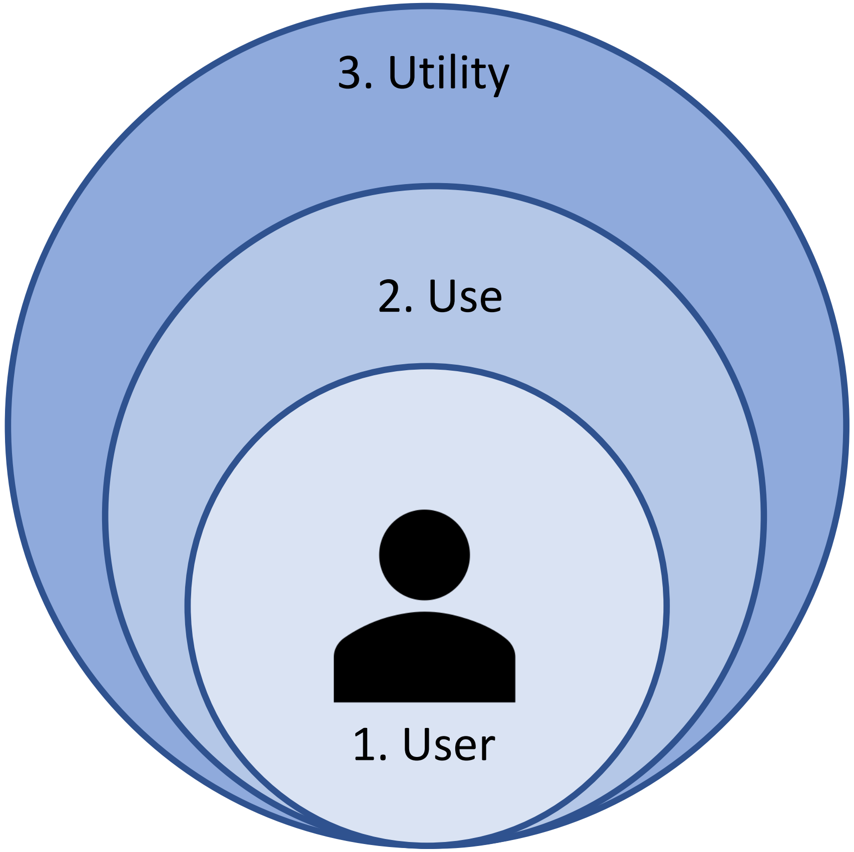
Effect of **Test-driven modeling** on **perceived cognitive load, perceived quality**, and **quality of the adapted models** in the context of maintainability tasks.

Results: The adoption of test cases

- could significantly reduce the perceived cognitive load
- could significantly improve the perceived quality
- while no significant effects on model quality could be shown

Source: Zugal et al. 2013

Evaluation Framework for DSR

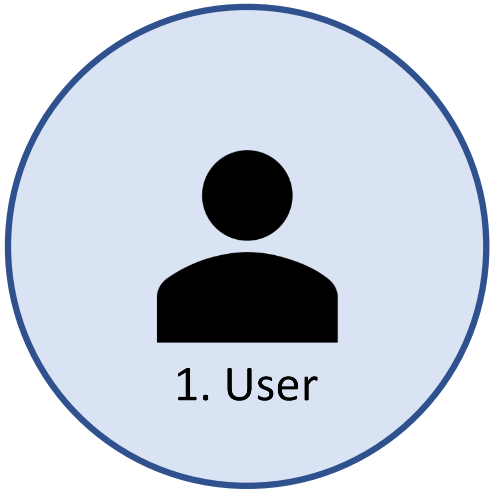


Utility emerges through the use of the artifact and depends on the user and the environment

Utility is relative!

Source: Mettler et al. 2014

Evaluation Framework for DSR



Study Subjects
Experimental Setting

Source: Mettler et al. 2014

Evaluation Framework for DSR: User

- **User** as the centerpiece of the evaluation framework
- The user determines **how** an artifact is **used** and what **value** she can gain from it

> ***Example**: Experiment to investigate the impact of two distinct business process modeling notations on model comprehension.*

Evaluation Framework for DSR: User

- **User** as the centerpiece of the evaluation framework
- The user determines **how** an artifact is **used** and what **value** she can gain from it

***Example**: Experiment to investigate the impact of two distinct business process modeling notations on model comprehension.*

Evaluation Framework for DSR: User

- **User** as the centerpiece of the evaluation framework
- The user determines **how** an artifact is **used** and what **value** she can gain from it

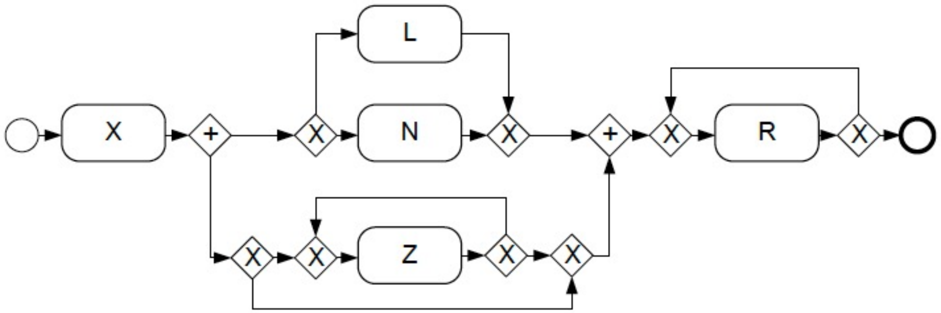
Example*: Experiment to investigate the impact of two distinct business process modeling notations on model comprehension.*

Study Subjects and Experimental Setting*:*

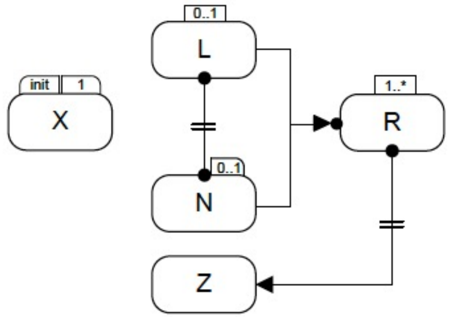
- Are study subject similarly knowledgeable in both modeling notations ?

Example: Imperative vs. Declarative Process Models

Controlled experiment comparing two modeling notations and two different task types in terms of accuracy and modeling speed.



Example of an imperative model



Example of a declarative model

Rather familiar

Limited familiarity

Source: Pichler et al. 2011

Evaluation Framework for DSR: User

- **User** as the centerpiece of the evaluation framework
- The user determines **how** an artifact is **used** and what **value** she can gain from it

***Example**: Experiment to investigate the impact of a particular method on modeling performance of master students.*

Evaluation Framework for DSR: User

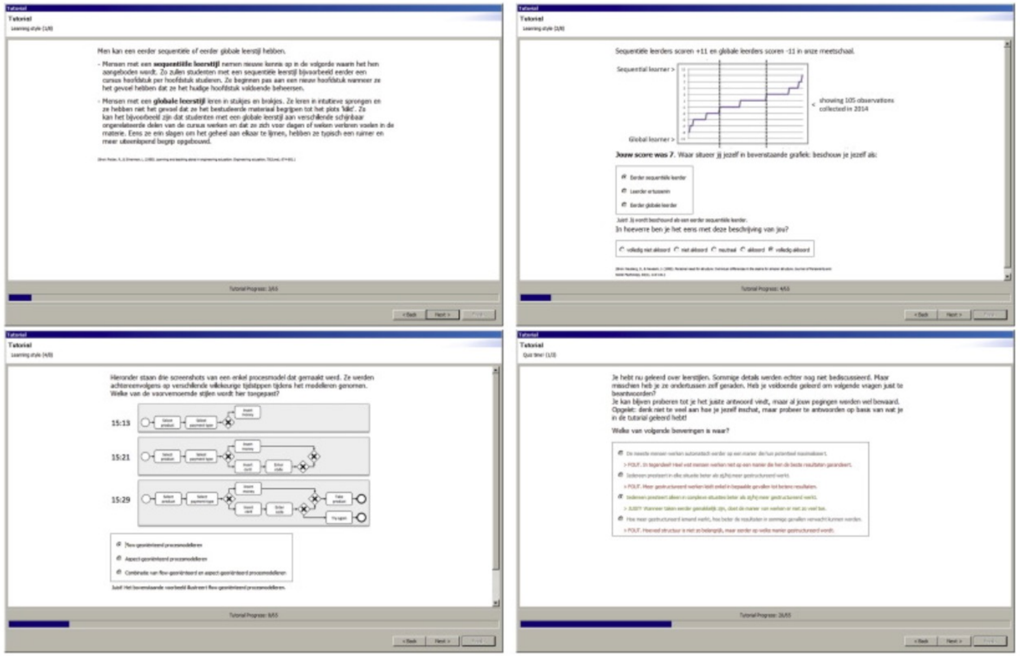
- **User** as the centerpiece of the evaluation framework
- The user determines **how** an artifact is **used** and what **value** she can gain from it

***Example**: Experiment to investigate the impact of a particular method on modeling performance of master students.*

***Study Subjects and Experimental Setting**:*

- Are study participants representative for the group of user for which the artifact was developed ?

Example: Structured Process Modeling Method



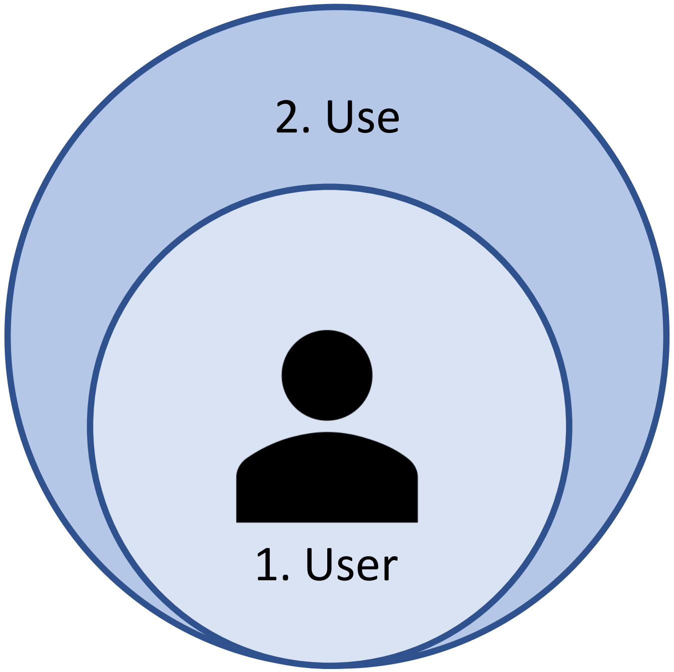
Method developed to support **modelers**.

146 **master students** of the Business Engineering program at Ghent University (Belgium) participated in the experiment

Paper **justifies** why master students have been chosen (instead of modeling practitioners or younger students)

Source: Claes et al. 2017

Evaluation Framework for DSR



Goals and scope of usage
Artifact characteristics
Manipulation procedure

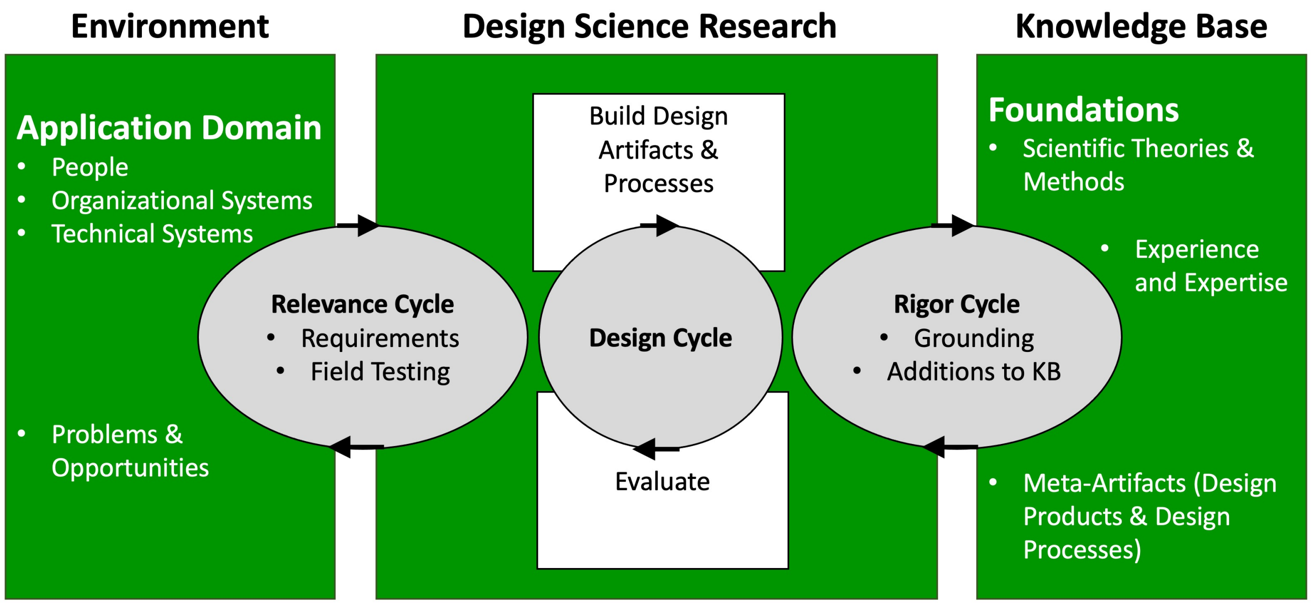
Study Subjects
Experimental Setting

Source: Mettler et al. 2014

Evaluation Framework for DSR: Use

- The **use** of the artifact is another crucial piece of information, since the situation on how the artifact is used influences its utility.

Scope and design goal for the developed artifact



Goal Type I

- Demonstrate that totally new artifact works

Goal Type II

- Demonstrate that developed artifact is better than existing ones

Evaluation Framework for DSR: Use

Example: *"A study in the area of business intelligence (BI) identifies a major problem in the representation of data. A prototype was developed in order to provide new means for visualizing data. Part-time MBA students were asked to perform distinct predefined tasks with the aim of comparing a traditional visualization with the newly developed representation. A questionnaire was used to capture the participants' personal beliefs on the usability of the solution."*

Source: Mettler et al. 2014

Evaluation Framework for DSR: Use

Example*: "A study in the area of business intelligence (BI) identifies a major problem in the representation of data. A prototype was developed in order to provide new means for visualizing data. Part-time MBA students were asked to perform distinct predefined tasks with the aim of comparing a traditional visualization with the newly developed representation. A questionnaire was used to capture the participants' personal beliefs on the usability of the solution."*

Scope and design goals are often reported in insufficient detail:

- For which type of data, which type of user, which domain, and which design goal (e.g., cost, quality, efficiency) was the artifact developed?

Evaluation Framework for DSR: Use

***Example**: "With the aim of improving the learning process of software developers, a novel method was designed which integrates additional information into an existing programming environment. The utility of the artifact is measured by a couple of metrics, such as the number of correctly answered questions related to a defined problem, the amount of time for finding deficiencies in programming code, or the total number of found deficiencies. The necessary data for evaluating the method was obtained from a questionnaire consisting of multiple-choice questions related to practical programming problems which was answered by undergraduate students."*

Source: Mettler et al. 2014

Evaluation Framework for DSR: Use

***Example**: "With the aim of improving the learning process of software developers, a novel method was designed which integrates additional information into an existing programming environment. The utility of the artifact is measured by a couple of metrics, such as the number of correctly answered questions related to a defined problem, the amount of time for finding deficiencies in programming code, or the total number of found deficiencies. The necessary data for evaluating the method was obtained from a questionnaire consisting of multiple-choice questions related to practical programming problems which was answered by undergraduate students."*

Artifact characteristics and manipulation procedure:

- Details concerning artifact characteristics and manipulation procedure is missing (**who does what, when, where, and how**)
- What programming problems were asked? How were the questions asked? In which situations were the students allowed to use the new method? What other auxiliary materials did the students have?

Source: Mettler et al. 2014

Evaluation Framework for DSR: Use

Overview of empirical studies into hierarchical structuring

Work	**Findings**
Moody [15] Domain: ER-Models	Positive influence on accuracy, no influence / negative influence on time
Reijers et al. [16,17] Domain: Business Process Models	Positive influence on understandability for one out of two models
Cruz-Lemus et al. [9,18] Domain: UML Statecharts	Series of experiments, positive influence on understandability in last experiment
Cruz-Lemus et al. [13] Domain: UML Statecharts	Hierarchy depth of statecharts has no influence
Shoval et al. [14] Domain: ER-Models	Hierarchy has no influence
Cruz-Lemus et al. [8] Domain: UML Statecharts	Positive influence on understandability for first experiment, negative influence in replication
Cruz-Lemus et al. [12,19] Domain: UML Statecharts	Hierarchy depth has a negative influence

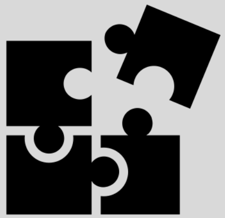
Source: Zugal et al. 2011

Evaluation Framework for DSR: Use

Overview of Empirical studies into hierarchical structuring

Work	**Findings**
Moody [15] Domain: ER	Positive influence on accuracy, no influence / neg-
Reijers et al. Domain: Bus	y for one out
Cruz-Lemus Domain: UM	ce on under-
Cruz-Lemus Domain: UM	influence
Shoval et al. Domain: ER	
Cruz-Lemus Domain: UML Statecharts	ity for first experiment, negative influence in replication
Cruz-Lemus et al. [12,19] Domain: UML Statecharts	Hierarchy depth has a negative influence

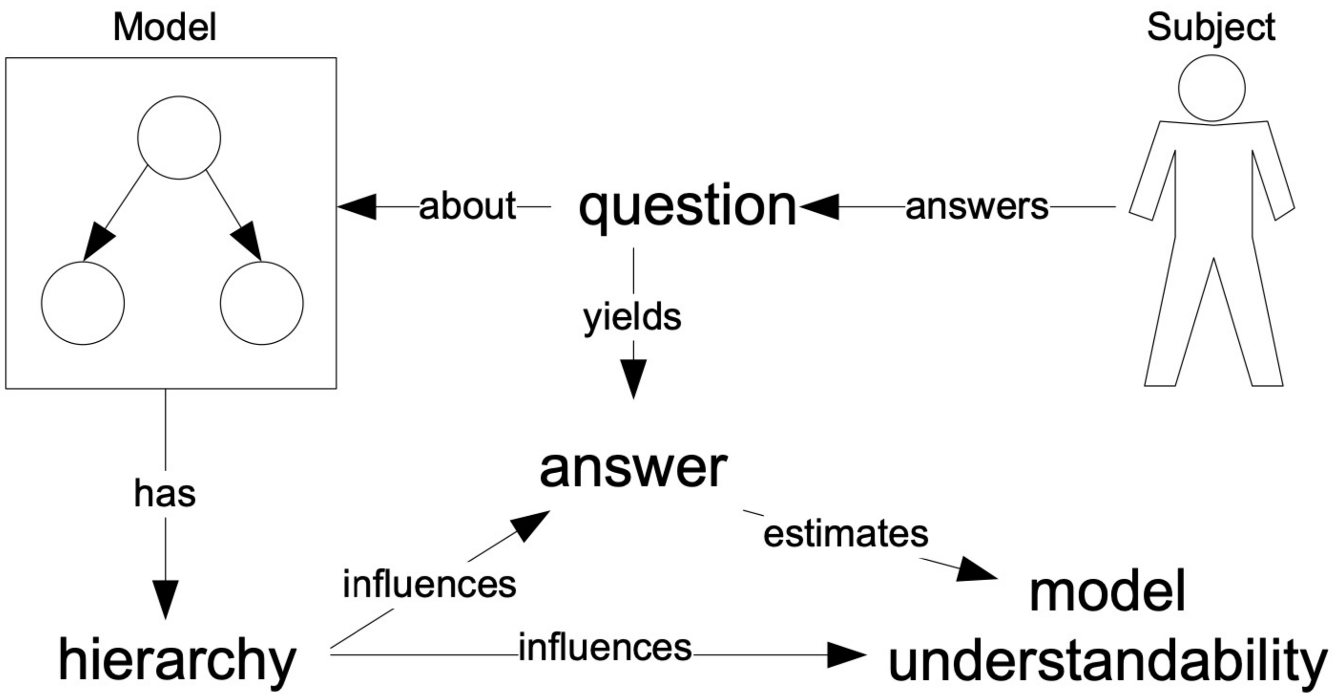
Inconclusive findings in literature



Artifact use as the missing piece

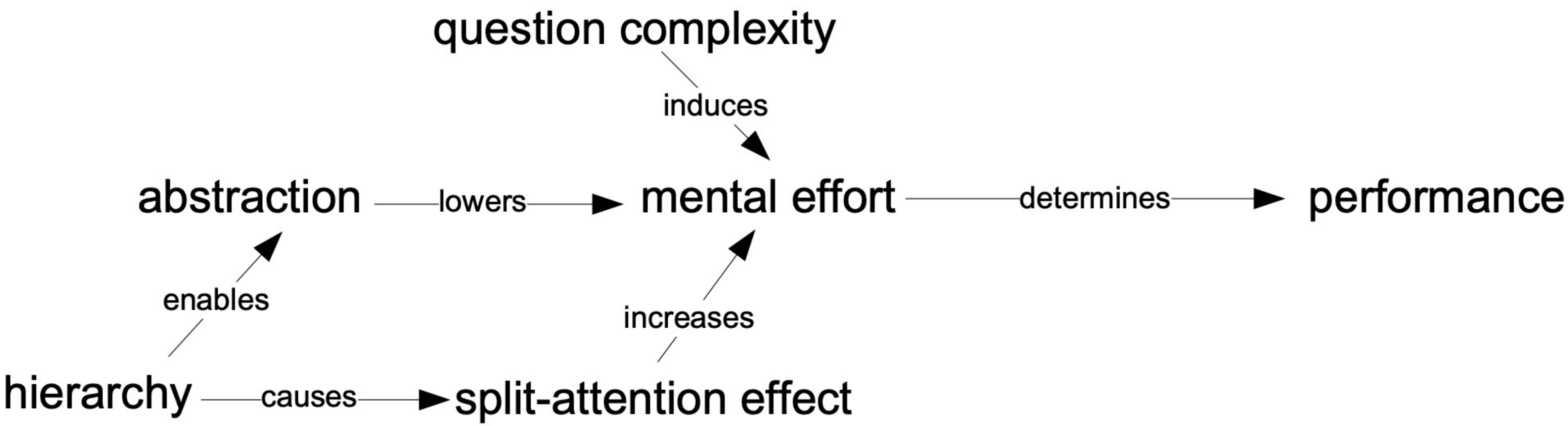
Source: Zugal et al. 2011

The Importance of the Task



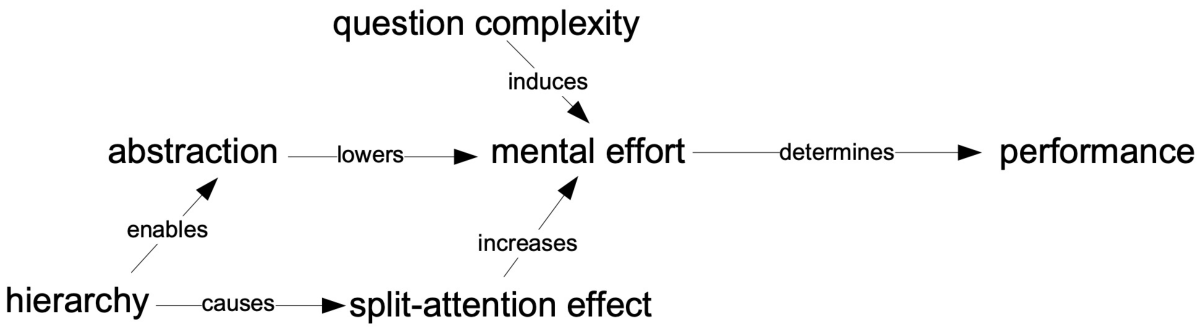
Source: Zugal et al. 2011

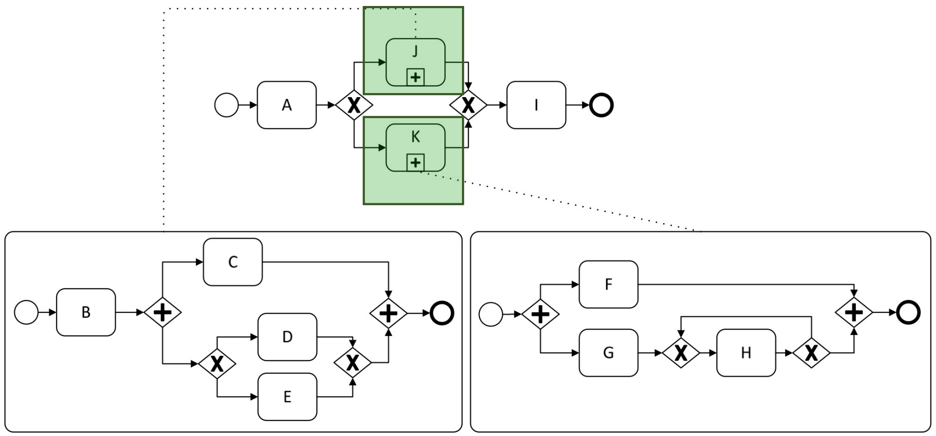
The Importance of the Task



Source: Zugal et al. 2011

The Importance of the Task

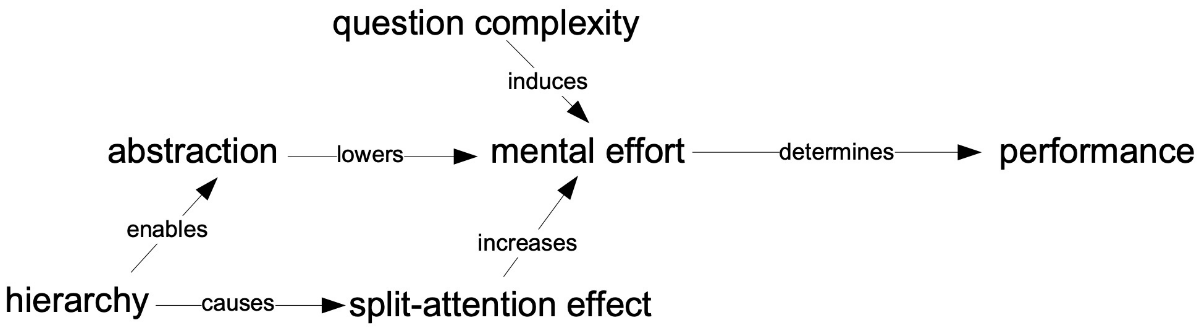


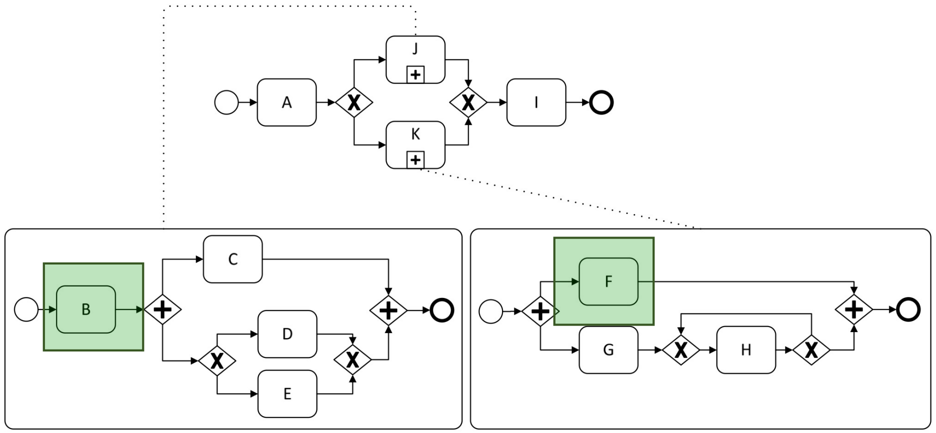


Are activities J and K mutually exclusive?

Source: Zugal 2013

The Importance of the Task

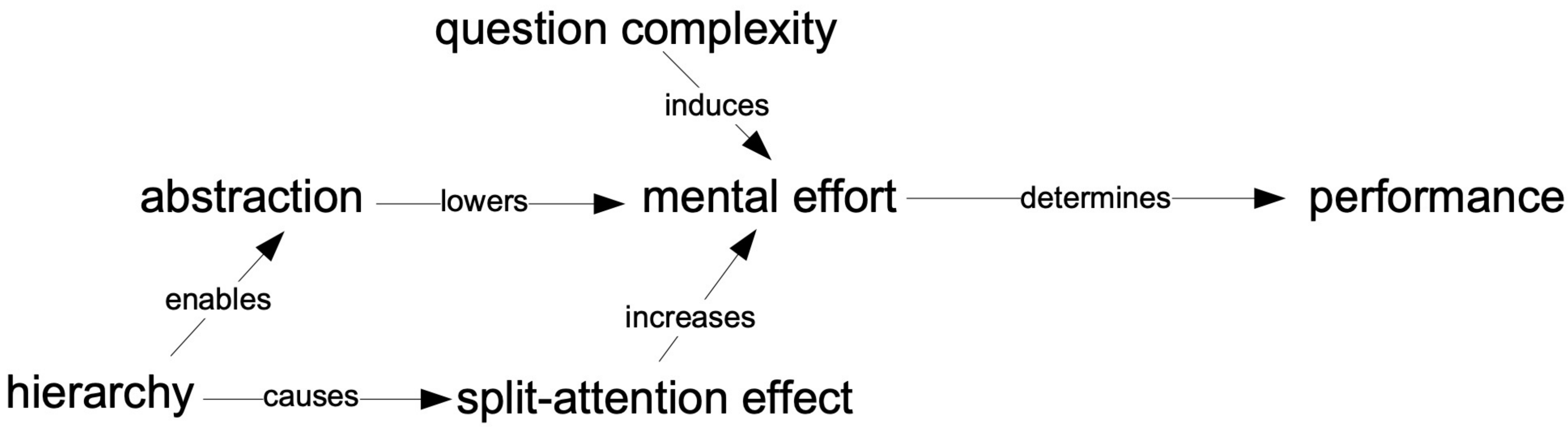




Are activities B and F mutually exclusive?

Source: Zugal 2013

The Importance of the Task



Key lesson: Results depend on *which* questions are asked or on a more abstract note on their *use*.

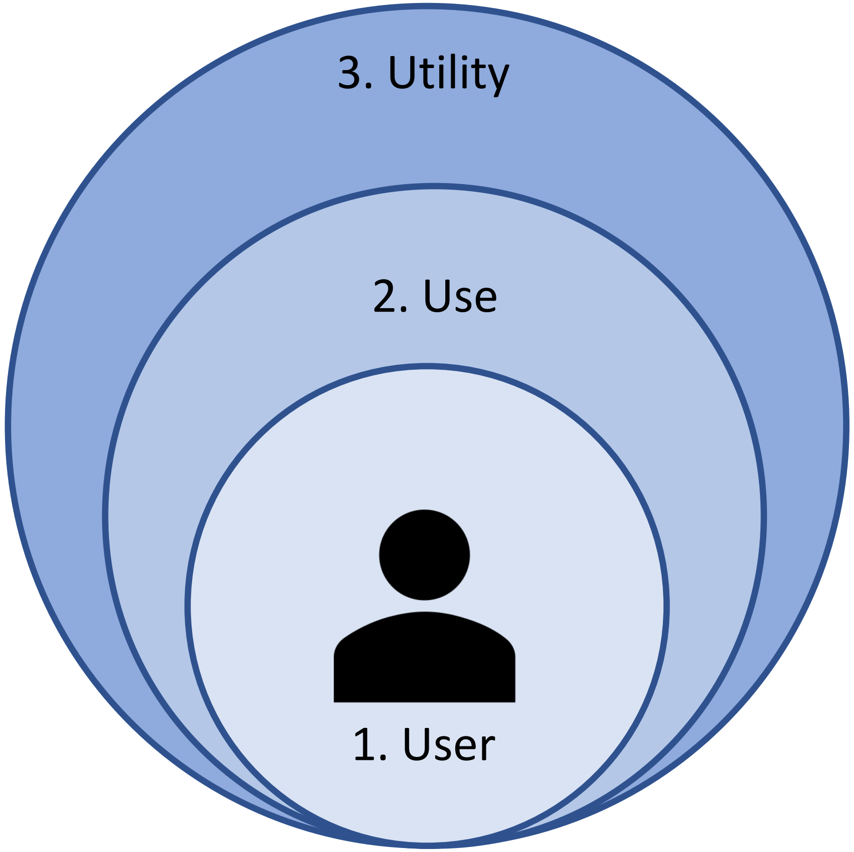
Source: Zugal et al. 2011, Zugal 2013

Evaluation Framework for DSR: Use

Key lesson: Results depend on *which* questions are asked or on a more abstract note on their *use*.

Precise description of **artifact characteristics** and **manipulation procedure** needed for replication. Ideally, a replication package can be provided.

Evaluation Framework for DSR



Evaluation metrics
Evaluation results

Goals and scope of usage
Artifact characteristics
Manipulation procedure

Study Subjects
Experimental Setting

Source: Mettler et al. 2014

Evaluation Framework for DSR: Utility

- **Utility** emerges through the use of the artifact and depends on the user and the environment
- **Utility is a relative concept**
- Clear and measurable variables are needed to assess utility
- Chosen metrics should give alternatives that are compared equal consideration (different alternatives might have been designed with different goals in mind)

Evaluation Framework for DSR: Utility

- Keep moderating and mediating effects in mind
  - User-specific characteristics (e.g., age, gender, and computer literacy)
  - Date and time (e.g., differences in bandwidth utilization depending on specific workdays)
  - Technical effects (e.g., divergent behavior of the designed artifact on different platforms)
  - Environmental effects (e.g., divergent behavior of the designed artifact due to temperature differences)
  - Socio-cultural effects (e.g., assignment of distinct connotations and meaning for the same artifact construct because of a different cultural background)

Evaluation Framework for DSR: Utility

Example: *"A study describes a new search algorithm for maximizing the proportion of useful hits. A design experiment was conducted with the aim to proof that the new algorithm provides more useful results than the hits of a commercial search engine. The "utility" was judged by means of user feedback. The metrics to measure search performance are "elapsed time for presenting search result" and "selectivity of responses," Metrics to describe the search quality are "number of good sources" (as defined by the user), "number of duplicates in results list," and "average list length."*

Source: Mettler et al. 2014

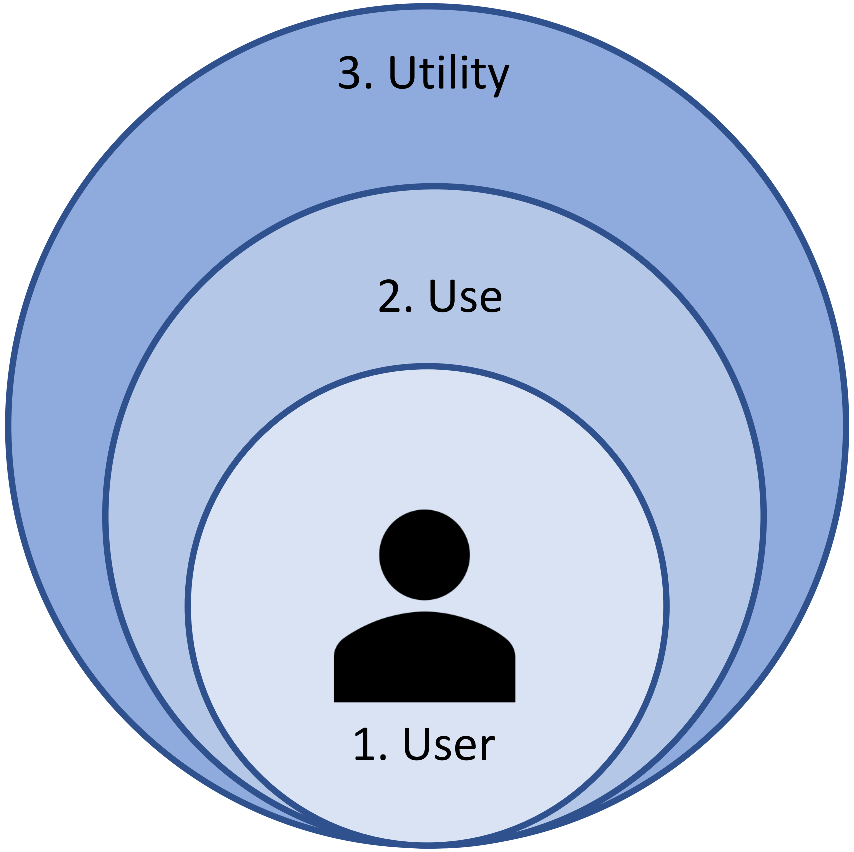
Evaluation Framework for DSR: Utility

***Example**: "A study describes a new search algorithm for maximizing the proportion of useful hits. A design experiment was conducted with the aim to proof that the new algorithm provides more useful results than the hits of a commercial search engine. The "utility" was judged by means of user feedback. The metrics to measure search performance are "elapsed time for presenting search result" and "selectivity of responses," Metrics to describe the search quality are "number of good sources" (as defined by the user), "number of duplicates in results list", and "average list length".*

- Importance of **mediating** and **moderating** factors:
  - Mediating and moderating factors play an important role in the scenario above (e.g., *goodness of hits might* be assessed differently by users depending on contextual and situational factors)

Source: Mettler et al. 2014

Evaluation Framework for DSR

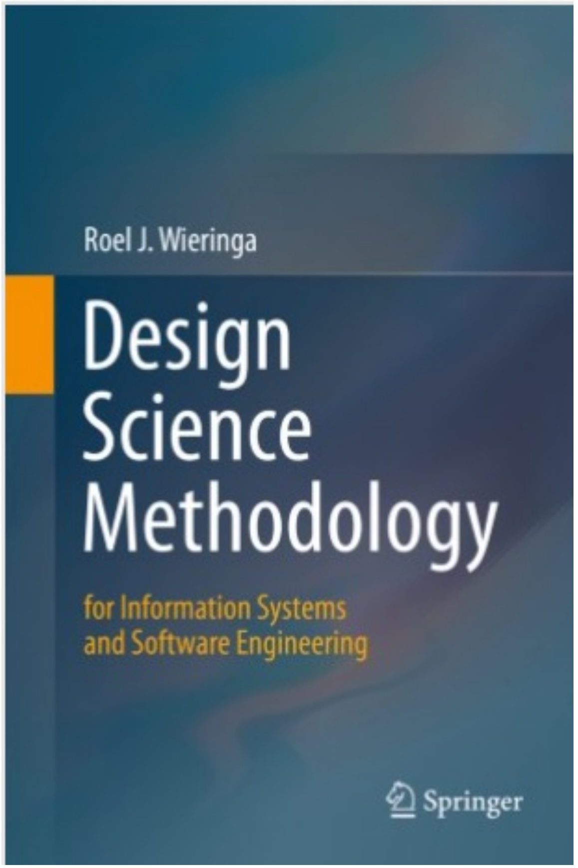


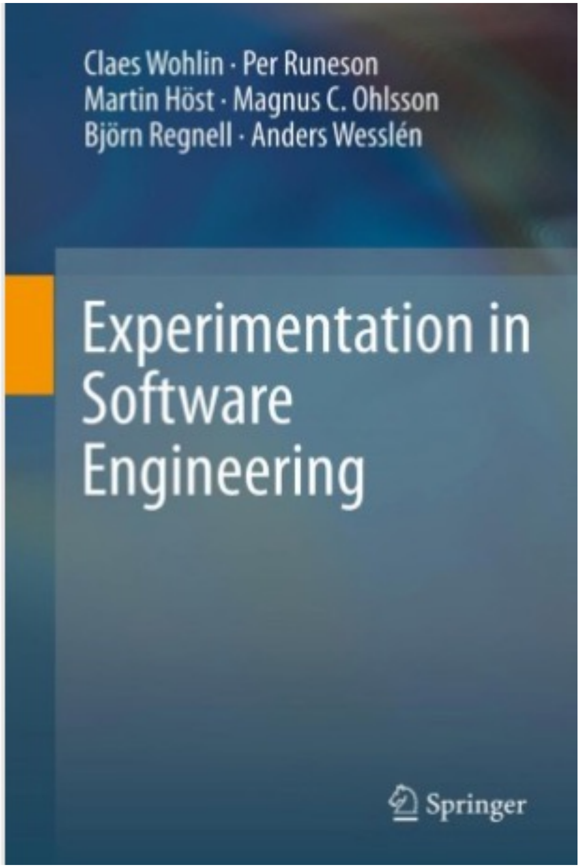
Utility emerges through the use of the artifact and depends on the user and the environment

Utility is relative!

Source: Mettler et al. 2014

Excellent Resources Available





References

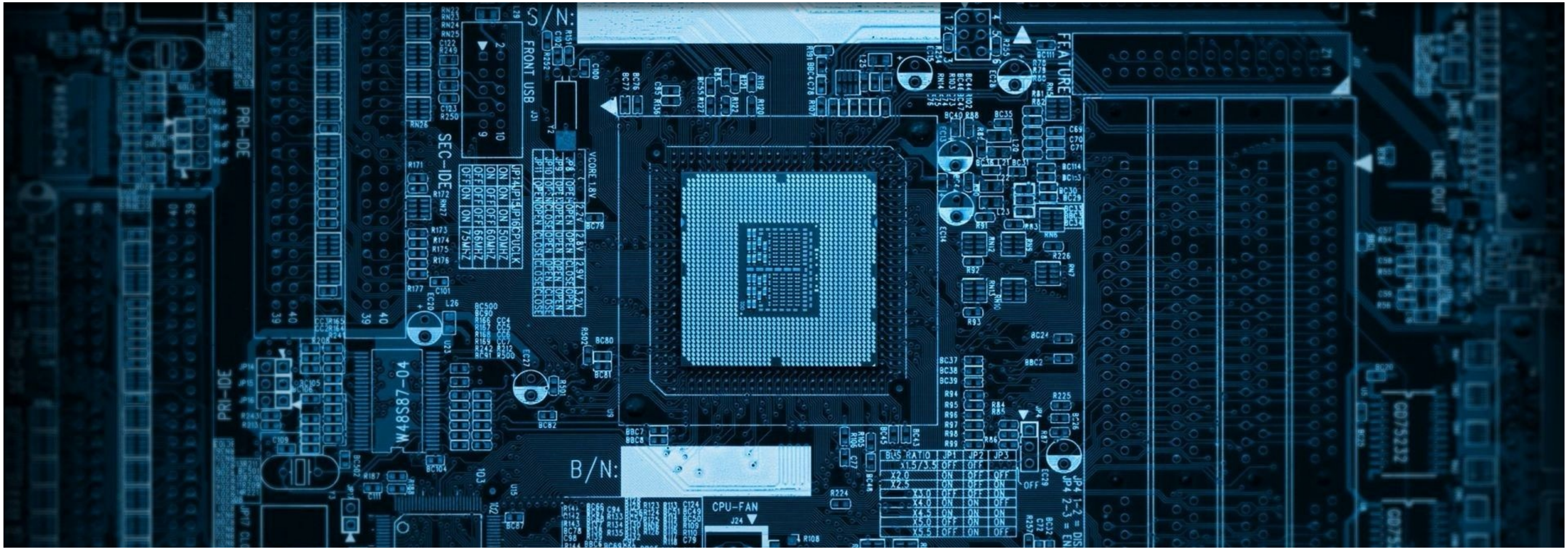
- *Anne Cleven, Philipp Gubler, Kai M. Hüner:*
  Design alternatives for the evaluation of design science research artifacts. *DESRIST 2009*
- *Alan R. Hevner, Salvatore T. March, Jinsoo Park, Sudha Ram:*
  Design Science in Information Systems Research. *MIS Q. 28(1): 75-105 (2004)*
- *Alan R. Hevner:*
  The Three Cycle View of Design Science. *Scand. J. Inf. Syst. 19(2): 4 (2007)*
- *Claes Wohlin, Per Runeson, Martin Höst, Magnus C. Ohlsson, Björn Regnell:*
  Experimentation in Software Engineering. *Springer 2012, ISBN 978-3-642-29043-5, pp. I-XXIII, 1-236*
- *Kathrin Figl, Jan Recker, Jan Mendling:*
  A study on the effects of routing symbol design on process model comprehension. *Decis. Support Syst. 54(2): 1104-1118 (2013)*
- *Tobias Mettler, Markus Eurich, Robert Winter:*
  On the Use of Experiments in Design Science Research: A Proposition of an Evaluation Framework. *CAIS 34: 10 (2014)*

References

- *Jan Claes, Irene T. P. Vanderfeesten, Frederik Gailly, Paul Grefen, Geert Poels:*
  The Structured Process Modeling Method (SPMM) what is the best way for me to construct a process model? *Decis. Support Syst. 100: 57-76 (2017)*
- *Paul Pichler, Barbara Weber, Stefan Zugal, Jakob Pinggera, Jan Mendling, Hajo A. Reijers:*
  Imperative versus Declarative Process Modeling Languages: An Empirical Investigation. *Business Process Management Workshops (1) 2011:* ***383-394***
- *Stefan Zugal, Jakob Pinggera, Barbara Weber, Jan Mendling, Hajo A. Reijers:*
  Assessing the Impact of Hierarchy on Model - A Cognitive Perspective. *EESSMod 2011*
- *Kevin Andrews, Sebastian Steinau, Manfred Reichert:*
  Enabling runtime flexibility in data-centric and data-driven process execution engines. *Inf. Syst. 101: 101447 (2021)*
- *Safdar Aqeel Safdar, Hong Lu, Tao Yue, Shaukat Ali, Kunming Nie:*
  A framework for automated multi-stage and multi-step product configuration of cyber-physical systems. *Softw. Syst. Model. 20(1): 211-265 (2021)*
- *Jan Recker, Michael Rosemann, John Krogstie:*
  Ontology- Versus Pattern-Based Evaluation of Process Modeling Languages: A Comparison. *Commun. Assoc. Inf. Syst. 20: 48 (2007)*

References

- *Andreas Lanz, Barbara Weber, Manfred Reichert:*
  Time patterns for process-aware information systems. *Requir. Eng. 19(2): 113-141 (2014)*
- *Clara Ayora, Victoria Torres, Barbara Weber, Manfred Reichert, Vicente Pelechano:*
  VIVACE: A framework for the systematic evaluation of variability support in process-aware information systems. *Inf. Softw. Technol.57: 248-276 (2015)*
- *Barbara Weber, Manfred Reichert, Stefanie Rinderle-Ma:*
  Change patterns and change support features - Enhancing flexibility in process-aware information systems. *Data Knowl. Eng. 66(3): 438-466 (2008)*
- *Nick Russell, Wil M. P. van der Aalst, Arthur H. M. ter Hofstede:*
  Workflow Patterns: The Definitive Guide. *MIT Press 2016, ISBN 9780262029827*
- *Yair Wand, Ron Weber:*
  On the ontological expressiveness of information systems analysis and design grammars. *Inf. Syst. J. 3(4): 217-237 (1993)*
- *Stefan Zugal, Cornelia Haisjackl, Jakob Pinggera, Barbara Weber :*
  Empirical Evaluation of Test Driven Modeling. *Int. J. Inf. Syst. Model. Des. 4(2): 23-43 (2013)*
- *Stefan Zugal:* ***Applying Cognitive Psychology for Improving the. Creation, Understanding and Maintenance of Business Process Models.*** *PhD thesis, Univ. of Innsbruck (2013).*


S/N:
FRONT USB
SEC-IDE
PRI-IDE
W48S87-04
B/N:
CPU-FAN
FEATURE
LINE OUT

Questions?