HyDEEP – WORKSHOP

Tokyo, IPA-SEC, December 14th, 2009

Hybrid Defect Content and Effectiveness Early Prediction

Presented by
Michael Kläs and Dr. Jens Heidrich

HyDEEP

Fraunhofer IESE

Welcome – About the Fraunhofer Gesellschaft

Named after

- Joseph von Fraunhofer (1787-1826), a successful researcher, inventor and entrepreneur

Role of the Fraunhofer Gesellschaft

- Germany's leading organization for applied research and technology transfer

Size

- 58 institutes
- Approx. 12.500 employees

Funding Volume

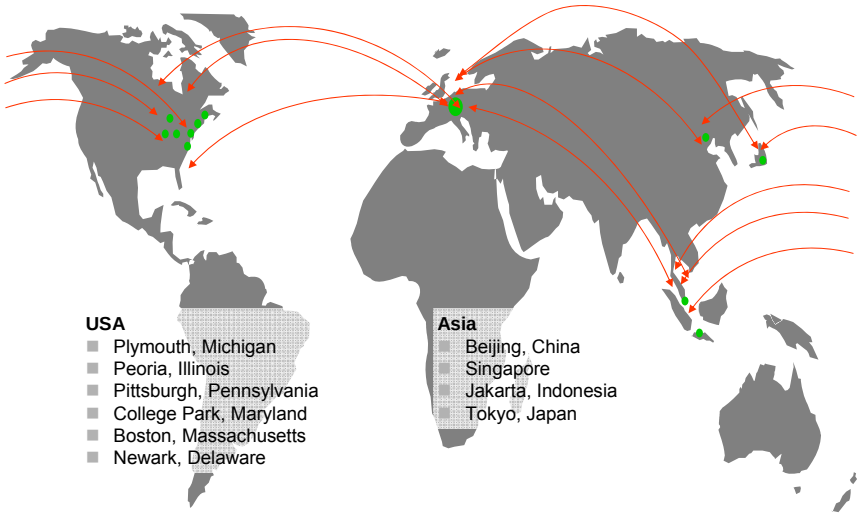
- about € 1.3 billion
- 1/3 base funding (government)
- 1/3 industrial projects
- 1/3 public sector projects

Fraunhofer IESE

Fraunhofer Research Units in Germany

- Institutes
- Branches of Institutes, Research Institutions, Working Groups, Branch Labs, and Application Centers

Fraunhofer Locations Worldwide



USA
- Plymouth, Michigan
- Peoria, Illinois
- Pittsburgh, Pennsylvania
- College Park, Maryland
- Boston, Massachusetts
- Newark, Delaware

Asia
- Beijing, China
- Singapore
- Jakarta, Indonesia
- Tokyo, Japan

About Fraunhofer IESE

Principles

- Transferring proven technologies into practice
- Applying empirical methods to evaluate processes and products
- Identifying improvement areas and proposing changes
- Utilizing experience to guide technical and management choices

Measurement Services

- Defining and optimizing quality assurance strategies
- Introducing and optimizing measurement systems
- Establishing and improving estimation capabilities
- Assessing products and processes
- Introducing measurement-based software process improvement
- Training and coaching

About the Presenters (1/2)

Michael Kläs

Researcher at the Processes and Measurement department (PAM) at the Fraunhofer Institute for Experimental Software Engineering (IESE), Kaiserslautern, Germany

Research focus

- Defect prediction & classification
- Software cost estimation
- Goal-oriented measurement
- Empirical software engineering

About the Presenters (2/2)

Dr. Jens Heidrich

Head of the Processes and Measurement department (PAM) at the Fraunhofer Institute for Experimental Software Engineering (IESE), Kaiserslautern, Germany

PAM focus on

- Goal-oriented Measurement
- Project Control Centers
- Domain-specific Quality Models
- Process Management and SPI

Short Introduction of Tutorial Participants

- Name
- Role in your organization
- Prior knowledge and experience
- Expectations regarding the tutorial

Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Overview: The HyDEEP Method

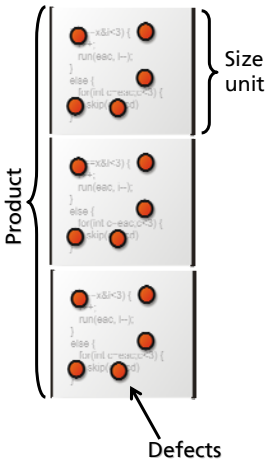
- Quality Management
- HyDEEP Application Possibilities
- HyDEEP Foundations
- HyDEEP Model Building Process

Overview: The HyDEEP Method

- Quality Management
  - Terminology and Challenges
- HyDEEP Application Possibilities
- HyDEEP Foundations
- HyDEEP Model Building Process

HyDEEP - Workshop
© Fraunhofer IESE
Fraunhofer IESE

Defect Density as a Software Quality Measure



Defect-based quality measures as a rough measure of overall software quality [Fenton97]

- defect density
  = number of known defects / product size unit

De-facto standard in industry

- due to limited measurement resources

Limitations

- No clear definition of what is a defect
- Not considering seriousness of a fault
- Difficult to measure size in a consistent and comparable way

How to handle these limitations

- Use only in one's own defined context
- Use formal, understood, and consistently applied definitions of defect and size
- Do not use for cross-organizational comparisons

Quality Assurance Activities

Quality assurance (QA) activities find and remove defects from a product/artifact.

Defect content: Initial number of defects when QA activity is performed

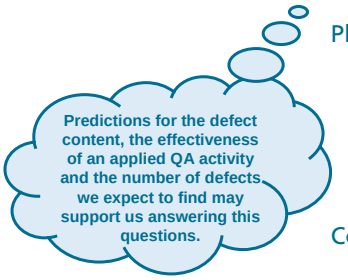
Defect found: Number of defects found and removed by performing the QA activity

Defects remaining / slippage: Number of defects remaining in the product

QA effectiveness (Eff) = defects found / defect content (here, e.g., 70%)

Remaining defect density (quality risk) = defect slippage / size

SW Quality Management – Critical Questions



Planning Product Quality

- How risky with respect to quality is my project compared to earlier projects?
- Which product quality can be expected?
- Do the quality assurance activities planned allow the achievement of quality objectives?

Controlling (Assuring) Product Quality

- How many defects do we have to remove in order to meet the project's quality objectives?
- Have quality assurance activities achieved the expected defect removal effectiveness?
- How many defects are expected to remain?

Improving Process Quality

- Which factors have the main impact on product quality in my project / organization?

Potential Answer – Predictive SW Quality Models

Capture-Recapture

Reliability Growth Models

COQUALMO

MARS for Inspections

Multivariate Regression Analysis

Many models

- address only some of the tasks/problems of quality management
- are applicable only to certain development phases (e.g., system test)
- do not allow stepwise introduction
- require large sets of high-quality measurement data for building custom-specific models or do not fit the specific context of an organization in the case of ready-to-use models

Applicability of existing predictive SW quality models in a specific context is typically limited

The HyDEEP Approach

HyDEEP addresses these problems by combining expert judgment and available measurement data to provide custom-specific guidance for managing software quality

- Support for planning, controlling, and improving quality assurance activities

HyDEEP transfers the well-tested basic principles of the cost estimation method CoBRA® to the quality assurance context

- Quantified causal models
- Monte Carlo simulation

Overview: The HyDEEP Method

- **Quality Management**
- **HyDEEP Application Possibilities**
  - Planning, Controlling, and Improving QA
- **HyDEEP Foundations**
- **HyDEEP Model Building Process**

Possible Application Purposes

P1: Identification of improvement potential

- Which factors have the main impact on quality?

P2: Early analysis of quality-related risks

- How risky with respect to quality is my project compared to earlier projects?
- Are the planned QA activities appropriate/justified?

P3: Prediction of defects found by QA activities

- How many defects do we have to remove in order to meet the quality objectives?
- Have we found the expected number of defects?

P4: Planning and managing quality assurance activities

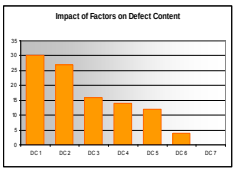
- How effective are the planned QA activities?
- How many defects are expected to remain?

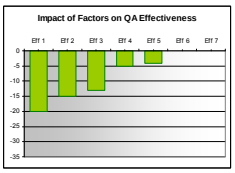
P1: Identification of Improvement Potential

Which factors have the main impact on quality?

How to answer?

- Use the quantified causal model to determine defect density and/or effectiveness increase caused by each factor
- Analysis is possible for actual projects or a group of historical projects in an organization





Identification of the factors with the highest impact on defect content and effectiveness

- Factors DC1 and DC2 have the highest impact on defect content
- Factor Eff1 offers the highest potential of increasing QA effectiveness

P2: Early Analysis of Quality-Related Risks

Low Risk

High Risk

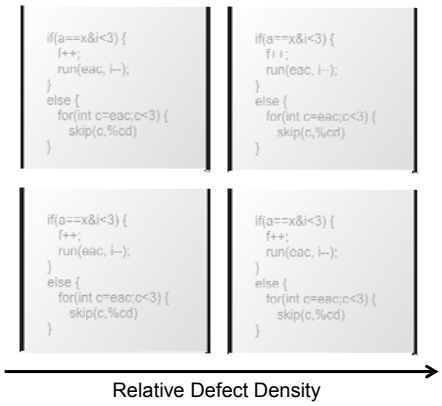
How risky with respect to quality is my project compared to earlier projects?

Are the planned QA activities appropriate/justified?

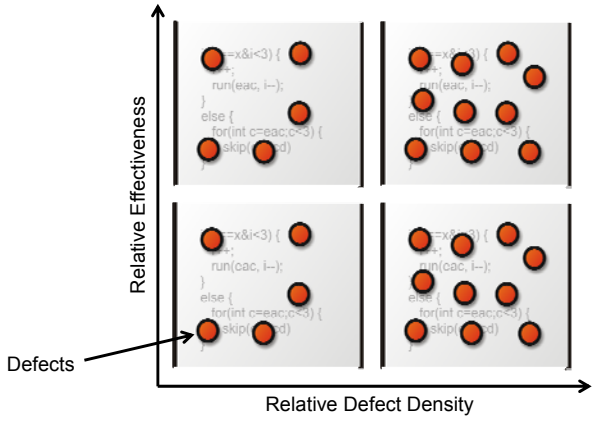
How to answer?

- Determine relative defect density and effectiveness of QA activity of the current project
- Compare with relative defect density and effectiveness historical projects

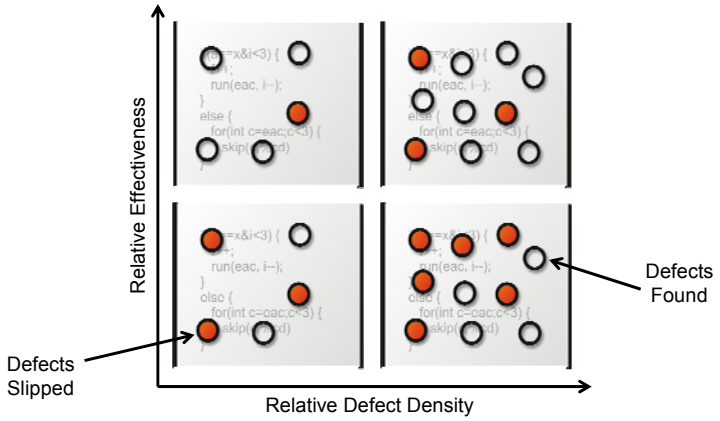
Planning Based on Relative Effectiveness/Defect Density



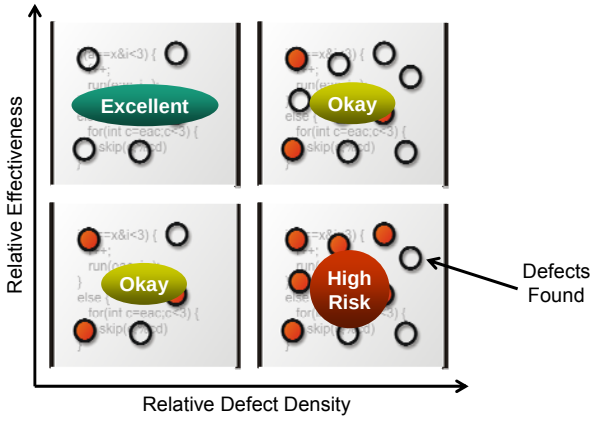
Planning Based on Relative Effectiveness/Defect Density



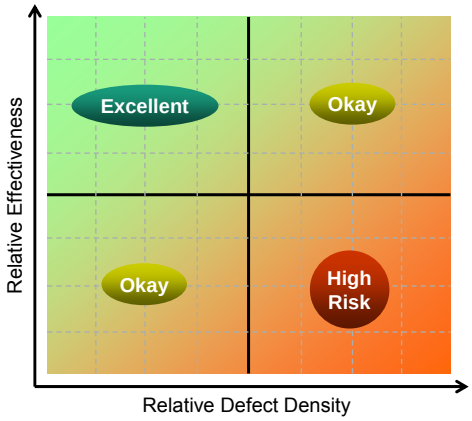
Planning Based on Relative Effectiveness/Defect Density



Planning Based on Relative Effectiveness/Defect Density



Planning Based on Relative Effectiveness/Defect Density

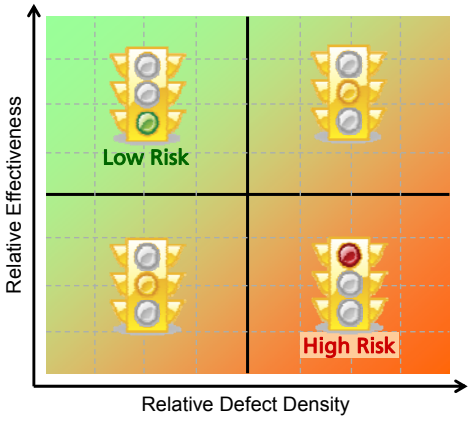


Planning Based on Relative Effectiveness/Defect Density

Low defect density and **high QA effectiveness**
Very low quality risk, but can also mean an inappropriately high QA intensity with respect to the defect density expected.

Low defect density and **low QA effectiveness**
No major quality risk - the relative number of defects that potentially slip through the QA activities is low.

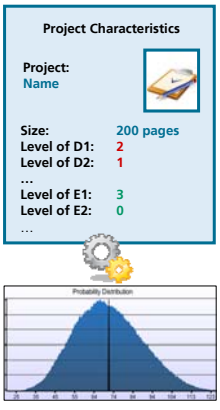
Quality Risk Analysis Chart



High defect density and **high QA effectiveness**
No major quality risk - the relative number of defects that potentially slip through the QA activity is low due to the high QA effectiveness

High defect density and **low QA effectiveness**
Major quality risk, since a relatively high number of defects can potentially slip through the QA activity and result in low quality of the delivered product.

P3: Quantitative QA Controlling



Number of defects expected to be found

How many defects do we have to remove in order to meet the quality objectives?

Have we found the expected number of defects?

How to answer?

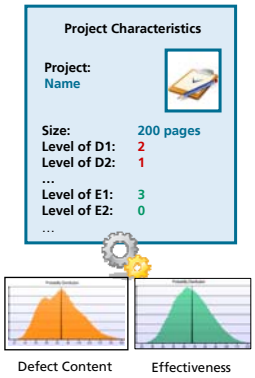
- Characterize the project with respect to
  - Defect content and effectiveness factors
  - Size of the artifact
- Predict the number of defects expected to be found
  - Based on quantitative causal model and
  - Number of defects found in earlier projects

Compare the number of defects found with the number of defects predicted to be found by the HDCE model

If the number of defects found is unusually high or low when compared with the predicted number of defects

- Check for incorrect factor levels or potentially missing relevant factors influencing effectiveness or defect content
- Re-estimate the remaining quality risk and adjust the QA activities (if needed)

P4: Quantitative QA Planning



How effective are the planned QA activities?

How many defects are expected to remain?

How to answer?

- Characterize the project with respect to
  - Defect content and effectiveness factors
  - Size of the artifact
- Predict defect content and QA effectiveness
  - Based on quantitative causal model and
  - Number of defects found by QA in earlier projects
  - Number of defects slipped QA in earlier projects

This information can be used to determine the expected number of remaining defects, e.g.,

- Expected Defect Content = 120
- Expected Effectiveness = 70%
- → Expected Remaining Defects = 36

Stepwise Introduction of HyDEEP

Data Requirements

ID: Purpose	Requirements*	Output
P1: QA Explanation / Improvement	Quantified causal model for QA activity; Characterization of actual project	Pareto chart identifying DC and Eff influencing factors in actual project with the highest improvement potential
P2: Qualitative QA Planning	(1) + size of checked artifact and characterization for ≥5-10 historical projects	Benchmarking of relative QA effectiveness and defect content of actual project against historical ones to identify projects with high quality risk
P3: Quantitative QA Controlling	(2) + number of defects found (DF) by QA activity for the historical projects	Thresholds for defects found by QA activity in actual project (based on DF probability distribution)
P4: Quantitative QA Planning	(3) + number of defects slipped (DS) through the QA activity for the historical projects	Prediction of absolute Eff and DC values for the actual project (i.e., actual DS can be predicted)

Method allows early benefits even if only few measurement data are available (see P1 & P2)

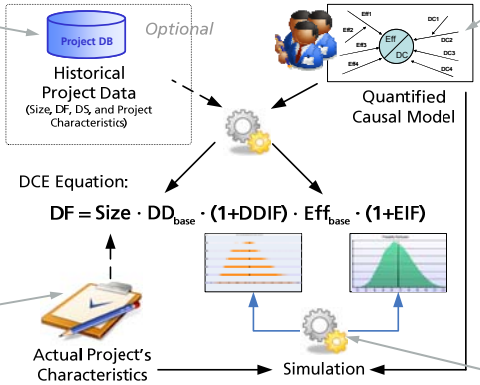
* Required measurement data are underlined.

Overview: The HyDEEP Method

- **Quality Management**
- **HyDEEP Application Possibilities**
- **HyDEEP Foundations**
- **HyDEEP Model Building Process**

HDCE* Model – Core of the HyDEEP Method

Historical project data are used to derive a defect content and effectiveness baseline for the context



A quantified causal model captures the expert opinions about relevant factors influencing defect content and effectiveness and their relative impact in the considered context

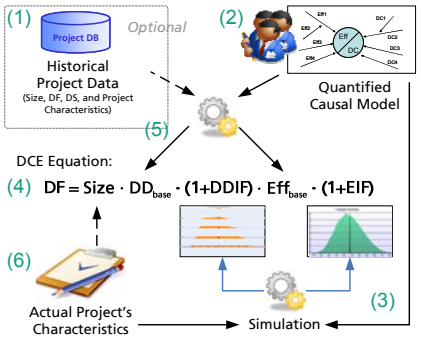
Expert-based characterization of the actual project allows determining the relative defect content and effectiveness probability distributions (relative to other projects in the context)

Monte Carlo simulation is applied to combine project characterization and the quantified causal model

* Hybrid Defect Content and Effectiveness

HDCE* Model: Illustration of the Model Components

1. Provide historical project data
2. Build a DCE casual model
3. Determine the project-specific deviations in defect density and QA effectiveness (by simulation)
4. Use the DCE equation
5. Determine the context-specific base values for defect density and QA effectiveness
6. Calculate prediction results for the current project



* Hybrid Defect Content and Effectiveness Model

Illustration of HyDEEP by an Midget Example

Context

- QA activity: Integration test
- Number of historical project
  - 5 historical releases of the product
- Limited measurement data
  - Externally developed components
  - No source code or other metrics

Application purpose

- Prediction of the expected number of defects to be found in the system test to control its performance and schedule times for defect fixing

Limitations of this example

- Hypothetical, strongly simplified causal model
- Only 2 influencing factors
  - typically 6 to 12 influencing factors
- Only direct factors (i.e., no interactions)

(1) Usage of Historical Project Data

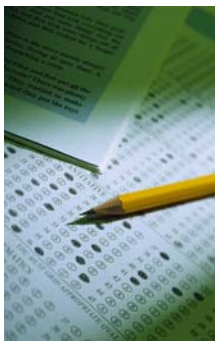
Measure for the change in a new release (size)

- Reason: Release with many/big changes contain usually more defects
- Possible measures
  - Delta Function Points
  - Delta Lines of Code
  - Number of new or changed features
- Chosen: theoretically required number of test cases to cover all changes in the product

Defect-related measures

- Number of defects found by integration test

(2) Strongly Simplified Causal Model with Two Factors

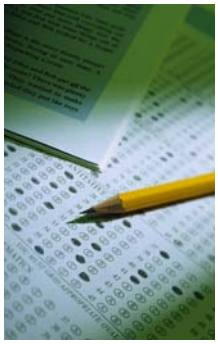


D1 Time pressure during development
(Impact on defect content)

- 0: No time pressure
- 1: Low time pressure
- 2: Increased time pressure
- 3: Extreme time pressure

Higher time pressure during development results in more defect in the tested product.

(2) Strongly Simplified Causal Model with Two Factors



E1 Testers inexperienced with the product
(Impact on QA effectiveness)

- 0: >75% of the test cases are conducted by testers with less then one year experience with the tested product
- 1: >50% (≤75%) of the test cases are conducted by testers with less than one year experience with the tested product
- 2: >25% (≤50%) of the test cases are conducted by testers with less than one year experience with the tested product
- 3: ≤25% of the test cases are conducted by testers with less than one year experience with the tested product

Tester with higher experience with the tested product find more defects during integration test resulting in a higher test effectiveness.

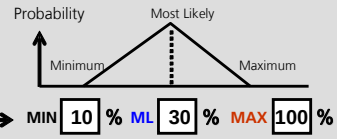
(2) Causal Model Quantification

Experts estimate for a factor the minimal, typical, and maximal magnitude of impact on the number of defect detected during integration testing, e.g.,

Factor D1: Time Pressure during development

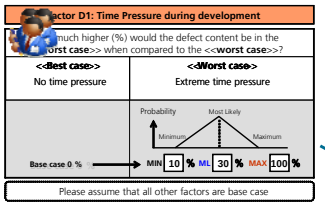
How much higher (%) would the defect content be in the <<**worst case**>> when compared to the <<**worst case**>>?

<<**Best case**>>	<<**Worst case**>>
No time pressure	Extreme time pressure

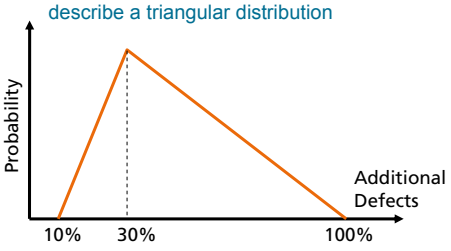


Please assume that all other factors are base case

(2) Causal Model Quantification



Estimated Min / most likely / max values describe a triangular distribution



E1 Inexp. Tester → Eff / DC ← D1 Time pressure

Table: impact of each factor is estimated by several experts

Expert	E1 Inexp. Testers			D1 Time pressure		
ID	Min	ML	MAX	Min	ML	MAX
1	30%	80%	120%	10%	30%	100%
2	20%	70%	90%	25%	40%	80%
3	60%	90%	120%	35%	60%	110%
4	50%	80%	150%	10%	30%	70%

(3) Determine Project-Specific Relative DD & Eff Increase

Expert	E1 Inexp. Testers			D1 Time pressure		
ID	Min	ML	MAX	Min	ML	MAX
1	30%	80%	120%	10%	30%	100%
2	20%	70%	90%	25%	40%	80%
3	60%	90%	120%	35%	60%	110%
4	50%	80%	150%	10%	30%	70%

(3/3) · ↓ (2/3) · ←

Expert	E1 Inexp. Testers			D1 Time pressure		
ID	Min	ML	MAX	Min	ML	MAX
1	30%	80%	120%	7%	20%	67%
2	20%	70%	90%	17%	27%	53%
3	60%	90%	120%	23%	40%	73%
4	50%	80%	150%	7%	20%	47%

Adapted project-specific table (project 1)

Project Characteristics

Project 1:
Release x.1

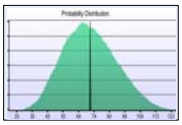
D1 Time Pressure
2: Increased time pressure

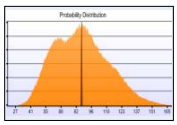
E1 Inexp. Testers
3: ≥25% of the test cases are conducted by testers with less then one year experience with the tested product

(3) Determine Project-Specific Relative DD & Eff Increase

Expert	E1 Inexp. Testers			D1 Time pressure		
ID	Min	ML	MAX	Min	ML	MAX
1	30%	80%	120%	7%	20%	67%
2	20%	70%	90%	17%	27%	53%
3	60%	90%	120%	23%	40%	73%
4	50%	80%	150%	7%	20%	47%

MC Simulation





EIF = E (probability distribution) = 80%

EIF = Effectiveness Improvement Factor

DDIF = E(probability distribution) = 33%

DDIF = Defect Density Increase Factor

E = Expectation value (i.e., average value of distribution)

(4) Defect Content and Effectiveness Equation

Effectiveness = Defects Found / Defect Content

⇕

Defects Found = Defect Content * Effectiveness

Defects Found = Size * Defect Density * Effectiveness

Defects Found = Size * DD_{base} * (1+DDIF) * Eff_{base} * (1+EIF)

(5) Determine Context-Specific Base Values

$(DD_{base} \cdot Eff_{base})_j$ can be determined for each historical project j = 1…n:

$$(DD_{base} \cdot Eff_{base})_j = DF_j / (Size_j \cdot (1+DDIF_j) \cdot (1+EIF_j))$$

- DDIF and EIF encapsulate the differences in DD and effectiveness between the projects
- DD_{base} and Eff_{base} values should be relatively stable in the context of the model
- Median of $(DD_{base} \cdot Eff_{base})_j$ can be used as a base value to estimate new projects in the context

$$(DD_{base} \cdot Eff_{base}) = Median\ ((DD_{base} \cdot Eff_{base})_{j=1..n})$$

Median is more robust against outlier then the average

(6) Predicting Expected Number of Defects Found

Project	Characterization	
Release	Factor E1	Factor D1
R1	3	2
R2	3	3
R3	1	3
R4	2	0
R5	0	1

Historical Projects

MC Simulation

Project	DF	EIF+1	DDIF+1	size	$(DD_{base} \cdot Eff_{base})_j$
Release	#Defects	Simul.	Simul.	# tTC	Result
R1	25	1,80	1,33	891	0,0117
R2	20	1,80	1,50	639	0,0116
R3	10	1,27	1,50	818	0,0064
R4	14	1,53	1,00	1620	0,0056
R5	5	1,00	1,17	651	0,0066

Median = 0,0066

Project	Characterization		size
Release	Factor E1	Factor D1	# tTC
R new	3	2	623

Current Projects

Defects Found =

$$Size \cdot (1+EIF) \cdot (1+DDIF) \cdot (DD_{base} \cdot Eff_{base})$$

$$623 \cdot (1+0{,}80) \cdot (1+0{,}33) \cdot 0{,}0066 \approx 10$$

MC Simulation

DDIF = Defect Density Increase Factor
EIF = Effectiveness Improvement Factor

Overview: The HyDEEP Method

- **Quality Management**
- **HyDEEP Application Possibilities**
- **HyDEEP Foundations**
- **HyDEEP Model Building Process**

HyDEEP Application Process – Major Steps

Initialize
- Motivate management
- Get commitment and budget

Characterize context
- Identify relevant environmental characteristics
- Identify stakeholders (champion, domain experts)
- Identify existing assets
  - Measurement data
  - HyDEEP models
- Identify constraints

Set goals
- Set up application objectives and scope
  - What's "in scope" vs "out of scope"
- Specify assumptions

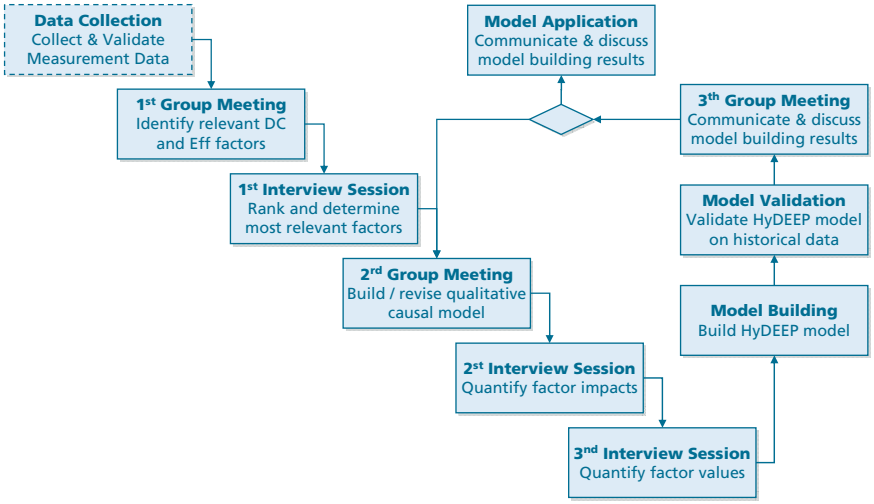
Plan application
- Plan resources
  - personnel & infrastructure
- Plan trainings
- Plan data collection and analysis
  - Measurement, group meetings, individual interviews
- Plan model development/adaptation

Apply HyDEEP
- Collect data
- Build/adapt and apply model

Validate application results
- Analyze model performance
- Identify improvement potentials

Package experiences
- Communicate results
- Package output assets (model, data, experiences)

Development of the HDCE Model – Example Scheduling



Development of the HDCE Model – Example Schedule (1/2)

Analysis of available measurement data

- Identification of the most relevant defect content and effectiveness (DCE) factors, their interactions, and their impacts

1st group meeting (factor selection)

- Brainstorming regarding potential DCE factors
- Ranking of identified factors with respect to their impact

Analysis of expert-based factor selection

- Analysis and aggregation of expert-based factor selection

2nd group meeting (qualitative causal model)

- Aggregation of expert- and data-based factor selection
- Quantification (scale definition) of selected DCE factors
- Expert-based identification of factor interactions
- Aggregation of expert- and data-based factor interactions

Development of the HDCE Model – Example Schedule (2/2)

Interview session (DCE multipliers and project data)

- Expert-based elicitation of DCE multipliers
- Expert-based elicitation of past project data

Analysis (model quantification)

- Analyze expert inputs with respect to completeness and consistency
- Analyze expert-based project data against the already elicited causal DCE model (the most relevant factors, factor interaction, and factor impact on DC or Eff)

3rd group meeting (quantitative model)

- Discuss results of most recent analysis and improve DCE causal model, if necessary
- Agree on the final model
- Select limited number of factors to be included in the DCE causal model

Scheduling can by slightly modified to fit existing time constraints

Activities with Expert Involvement: Initial Iteration

Activity	Purpose	#Experts	Effort per expert
1st Workshop	Identification of relevant factors and available data	3	~3.5h
1st Survey	Ranking of factors	4	~20 min
2nd Workshop	Discussion of ranking results, building of the causal model and introduction of the 2nd and 3rd survey	3	~1h
2nd Survey	Quantification of factor impact	4	~25min
3rd Survey	Collecting historical project data (i.e., quantify factor values)	3	~1h
3rd Workshop	Presentation and discussion of model and results (if required plan next iteration)	3	~2h
	Total	**3-4**	**≥ 1 day**

Activities with Expert Involvement: Initial Iteration

1st Workshop
- Status regarding Measurement/Defect data
- Identifying, collecting, and classifying relevant influencing factors

1st Survey: Ranking
- Evaluating influencing factors w.r.t. their relevancy

2nd Workshop
- Reporting survey results and factor selection
- Build qualitative DCE causal model
- Preparing quantification and data collection steps

2nd Survey: Impact Quantification
- Quantifying selected influencing factors

3rd Survey: Collecting project data
- Collecting data from historical projects (size, defects, influencing factors)

3rd Workshop: Feedback
- Presenting results of defect modeling
- Evaluating performance of the HyDEEP model

Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Characterize Context

Limit the HDCE model application context

- Start small: reduce the number of potential factors influencing defect content and QA effectiveness

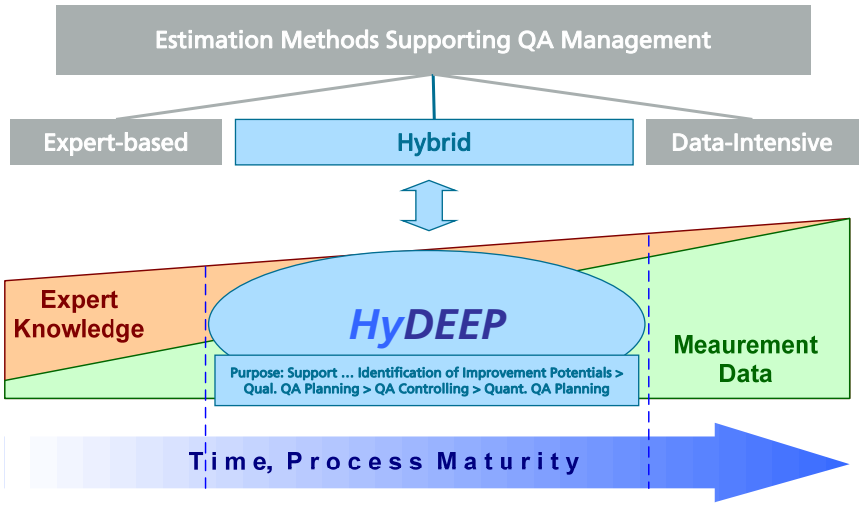
Characterize current processes

- Measurement and data collection
- Estimation approaches currently in use

Characterize capabilities and limitations

- Availability of domain experts
- Availability of quantitative data about QA applications (already completed)

Availability of Experience and Data



Essential Questions to Ask before We Start (1/2)

What is the scope of the model?

- Phase / QA activity:

What are my objectives?

- Process Improvement
- QA Planning
- QA Controlling

What is my context?

- Domain:
- Organization size:
- Technologies used:
- Artifacts / documents:
- New development or enhancement:
- Process maturity:

What are my assumptions?

- …

Essential Questions to Ask before We Start (2/2)

What measurement data are available?

- Size measure for the artifact:
- Defect found by QA activity:
- Defect sipping the QA activity:

How good is the data quality? (Reliability/validity)

What is the number of QA applications (already completed)?

- x applications / projects / releases

How is the QA (effort) planned?

- Is effort planned according to artifact size / size of change?
- Is there an existing estimation approach?

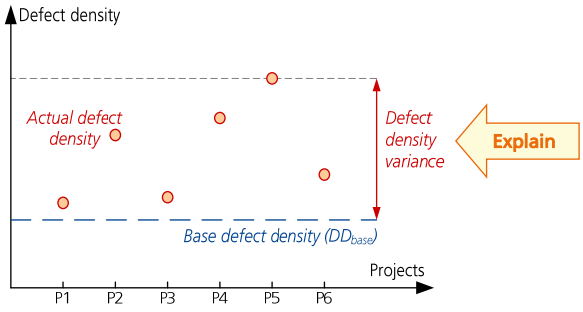
Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Basic Idea Underlying the HyDEEP Method

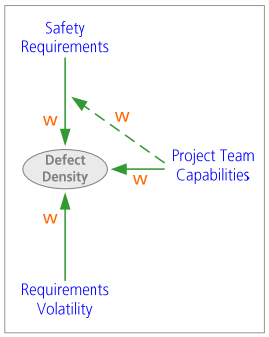
Build a DCE causal model that allows explaining in a certain context the variance over the project in

- the defect density of the quality assured product



Distribution of actual defect density

Explain

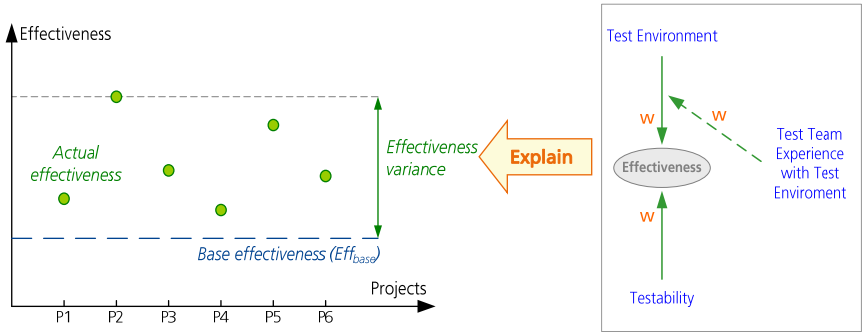


Causal defect density model

Basic Idea Underlying the HyDEEP Method

Build a DCE causal model that allows explaining in a certain context the variance over the project in

- …and the detection effectiveness of the QA activity



Distribution of actual effectiveness **Causal effectiveness model**

Initial Set of Defect Content and Effectiveness Factors

Perform brainstorming session

- Elicit factors that are relevant defect content and effectiveness drivers from the expert's perspective (experience)
- Relevant means
  - the level of the factor varies across projects in the considered context
  - the level of the factor can be determined or at least reasonably judged for each project
  - the experts assume that the variation has a noticeable impact on DC or Eff
- Obtain common agreement with regard to factor definitions (common understanding)
- Note for each factor
  - a concise factor name, a short definition, a category, and
  - a context-specific and realistic best case and worst case
- Check completeness with the reference list
  - ensure together with the experts that no relevant factor is missed

Basic Terminology of the HyDEEP Method

Base case describes a realistic situation in the context where the considered factor leads to a minimal increase in the number of defects found (per size unit)

Extreme case describes a realistic situation in the context where the considered factor results in a maximal increase in the number of defects found (per size unit)

- Note that base and extreme cases should not be the ones that can be imagined but those that are possible within a selected context (within which the HDCE model is built).

Best case is the "best" case that is still possible within a specified context.

- For instance, in a company X, the best case of a factor "Requirements volatility" means that less than 5% of requirements change after the requirements specification phase. In company Y, it might be less than 10%.

Worst case is the "worst" case that is still possible within a specified context.

Factor Type	Base Case	Extreme Case
Defect Content	Lowest → Best Case	Highest → Worst case
Effectiveness	Lowest → Worst Case	Highest → Best Case

Categories of Defect Content and Effectiveness Factors

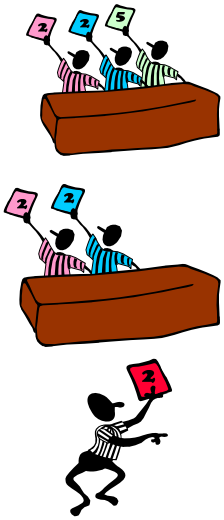
Defect Injection		Detection Factors	
1	Developer capability	1	Testability
2	Domain knowledge	2	Product complexity
3	Team composition	3	Quality of documentation
4	Team distribution	4	Change control
5	Collaboration	5	Test planning
6	Business management maturity	6	Management attitude
7	Product complexity	7	Adherence to plan
8	Communication	8	Test process maturity
9	Project management maturity	9	Development process maturity
10	External disturbances	10	Test environment
11	Process maturity	11	Support for testing
12	Change control	12	Product integration
13	Quality of documentation	13	Test capability
14	Requirements	14	Test team cohesion
15	Development environment	15	Team distribution
16	Innovation	16	Test team organization
		17	Communication

Jacobs J, van Moll J, Kusters R, Trienekens J, Brombacher A (2007) Identification of factors that influence defect injection and detection in development of software intensive products. Inf. Softw. Technol., vol. 49, no. 7, pp. 774-789

Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Ranking Defect Content and Effectiveness Factors



Perform ranking

- Perform separate ranking for each factor category
- Use sorting cards instead of weighting factors from the list in order not to suggest any specific order

Analyze ranking results

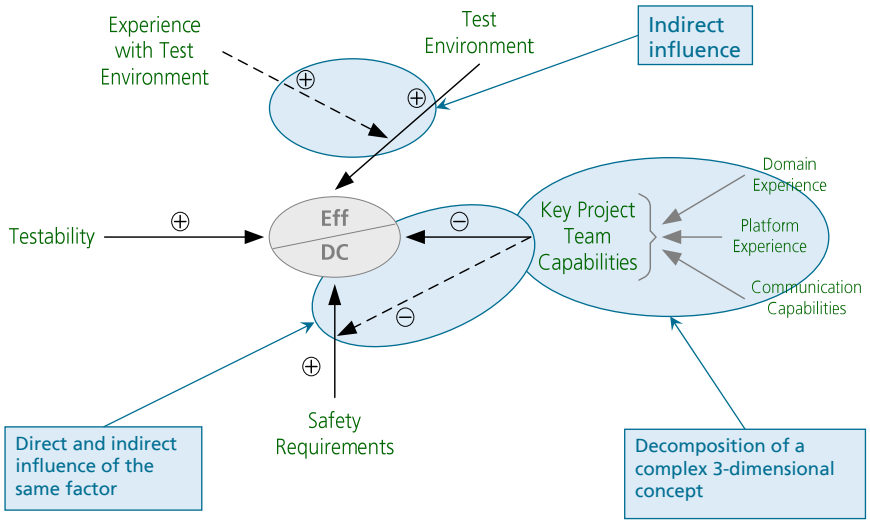
- Calculate Min, Max and Range for each factor to evaluate preference and consensus among experts
- Compute Kendall's coefficient of concordance
- Compare results when excluding outlying experts (e.g., the least experienced experts)

Aggregate ranking results and select the most significant factors

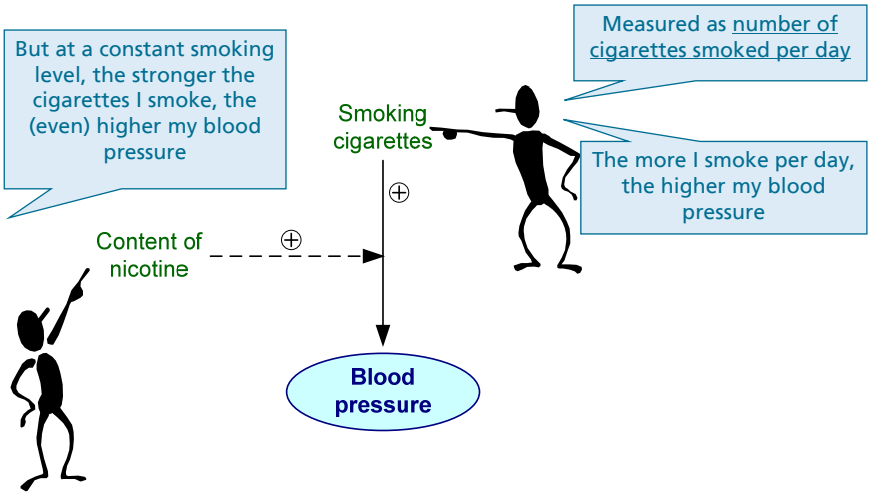
Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

STEP 2.3 – Types of Factor Interactions



STEP 2.3 – Direct vs. Indirect Interaction



Fraunhofer IESE

Identification of Factor Interactions

Identify direct interactions

- Are there any factors that influence the value of a certain factor?
- Which is easier to measure: cause or effect? Remove the one that is more difficult to measure.

Identify indirect interactions

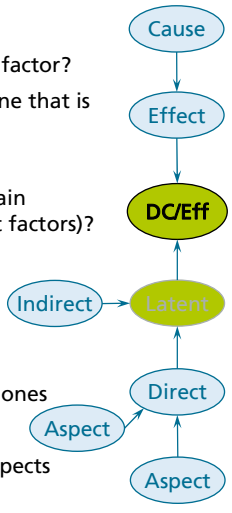
- Are there any factors that influence the strength of a certain factor's impact on defect content or effectiveness (indirect factors)?
- Consider only the most significant interaction (contributing to large variance of impact)

Agree on interactions

- Add new factors (to already selected ones), if necessary
- Reduce the number of interactions to the most important ones

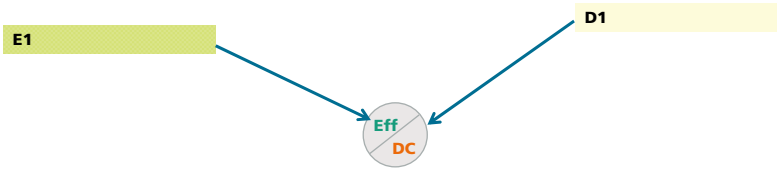
Decompose complex factors

- Review each factor's definition with respect to multiple aspects
- Consider only the most relevant aspects of the factor



Fraunhofer IESE

Causal Model (Template)



Workshop Agenda

Morning (AM)	10:00 11:00	Session I: HyDEEP Overview
	11:00 11:30	Session II: Scope Definition
	11:30 12:30	Session III: Identify Influencing Factors
		Lunch
Afternoon (PM)	13:30 14:00	Session III: Identify Influencing Factors (cont.)
	14:00 15:00	Session IV: Ranking Influencing Factors
	15:00 16:00	Session V: Causal Model Building & Next Steps

Activities with Expert Involvement: Initial Iteration

Activity	Purpose	#Experts	Effort per expert
1st Workshop	Identification of relevant factors and available data ✓	3	~3.5h
1st Survey	Ranking of factors ✓	4	~20 min
2nd Workshop	Discussion of ranking results, building of the causal model and introduction of the 2nd and 3rd survey ✓	3	~1h
2nd Survey	Quantification of factor impact	4	~25min
3rd Survey	Collecting historical project data (i.e., quantify factor values)	3	~1h
3rd Workshop	Presentation and discussion of model and results (if required plan next iteration)	3	~2h
	Total	**3-4**	**≥ 1 day**

Elicitation of Defect Content and Effectiveness Multipliers

Definition of multiplier

- Percentage increase relative to base case

Model expert's uncertainty

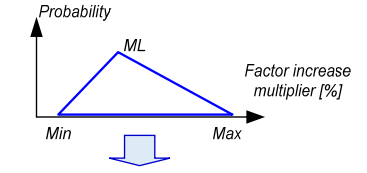
- Three values (triangular distribution): Min, Most Likely, Max increase

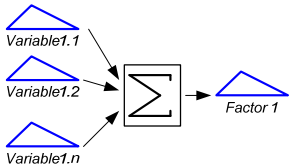
Direct vs. indirect influence

- Multipliers are estimated only for factors that directly contribute to DC or Eff increase.
- By indirect influence, base and extreme cases of an indirect factor are considered to quantify a direct factor's influence

Factor aspects (variables)

- For composite factors, only multipliers of related variables are assessed. The factor's multiplier is a sum of multipliers over its component variables.





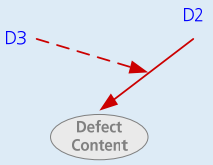
Defect Content Multipliers for Direct and Indirect Factors

Direct influence

D1: How much higher (%) would the defect content be in the <worst case> when compared to the <best case>?

Multiplier (increase relative to base case) in extreme case [%]	___ /___ /___ Min ML Max

Indirect influence



D2: How much higher (%) would the defect content be in the <worst case> when compared to the <best case>?

Multiplier (increase relative to base case) in extreme case [%] IF factor D3 is	
<Base>	<Extreme>
___ /___ /___ Min ML Max **M1**	___ /___ /___ Min ML Max **M2**

Fraunhofer IESE

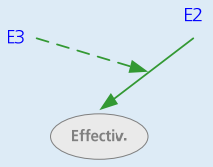
Effectiveness Multipliers for Direct and Indirect Factors

Direct influence

E1: How much higher (%) would the number of defects found be in the <best case> when compared to the <worst case>?

Multiplier (increase relative to base case) in extreme case [%]	___ /___ /___ Min ML Max

Indirect influence



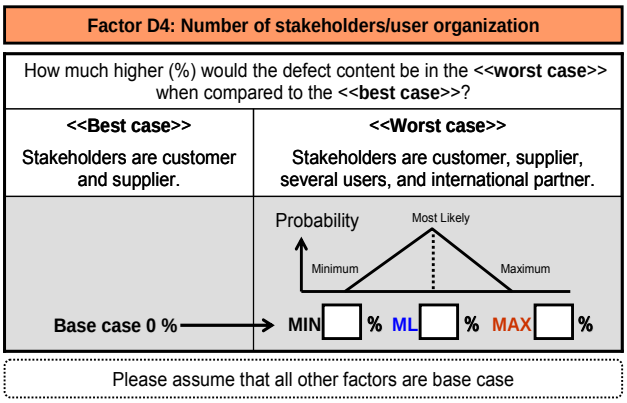
E2: How much higher (%) would the number of defects found be in the <best case> when compared to the <worst case>?

Multiplier (increase relative to base case) in extreme case [%] IF factor E3 is	
<Base>	<Extreme>
___ /___ /___ Min ML Max **M1**	___ /___ /___ Min ML Max **M2**

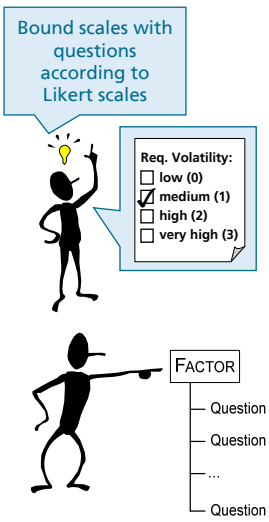
Fraunhofer IESE

Question Asking for Defect Content Factor's Impact

Questionnaire visually supporting the expert in factor impact quantification



Quantify Selected DCE Factors (Expert-based Factors)

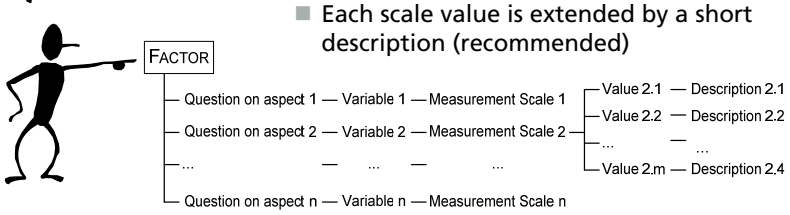


Define scales for identified DCE factors

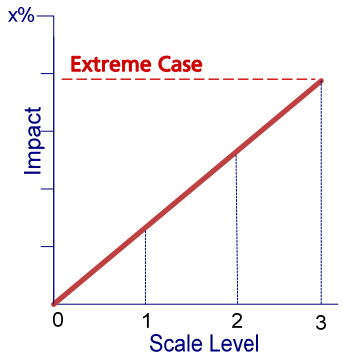
- define scales for factors together with one or more domain experts (if possible)
- obtain approval/agreement of involved domain experts (mandatory!)

General schema

- Scale is defined for each variable
- The 4-grade Likert scale is used
- Each scale value is extended by a short description (recommended)



Defect Content Factor with Scale-Level Description



D4: Number of stakeholder/user organizations	
Level	**Description**
0	Stakeholders are customer and supplier.
1	… customer, supplier, and user.
2	… customer, supplier, and international partner.
3	… customer, supplier, several users, and international partner.

Assumption of linear increase of factor impact over all defined scale levels

Thank you for your attention!

Michael Kläs
Phone: +49 (631) 6800 2110
Fax: +49 (631) 6800 1398
Email: michael.klaes@iese.fraunhofer.de

Dr. Jens Heidrich
Phone: +49 (631) 6800 2193
Fax: +49 (631) 6800 1398
Email: jens.heidrich@iese.fraunhofer.de