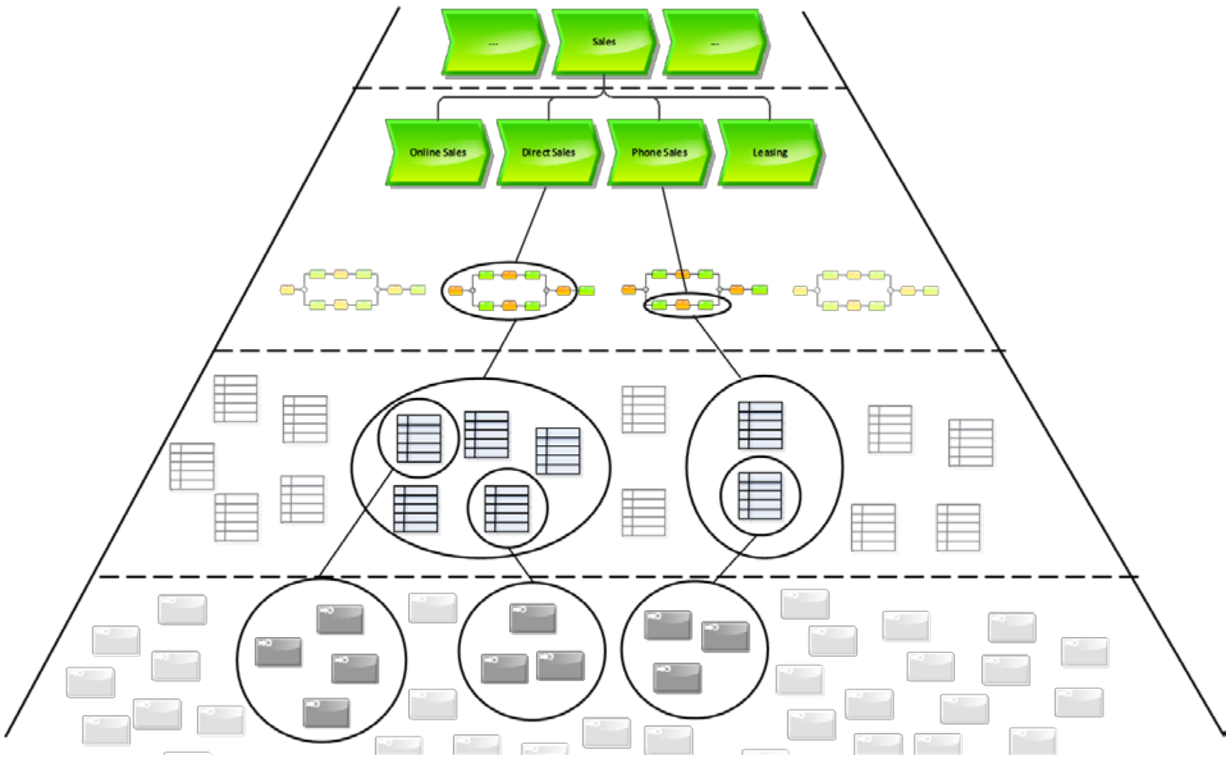
AN ANALYSIS OF PRIORITY-BASED DECISION HEURISTICS FOR OPTIMIZING ELICITATION EFFICIENCY

Norman Riegel, Jörg Dörr

{norman.riegel,joerg.doerr}@iese.fraunhofer.de



Fraunhofer IESE

A short quote from yesterday…

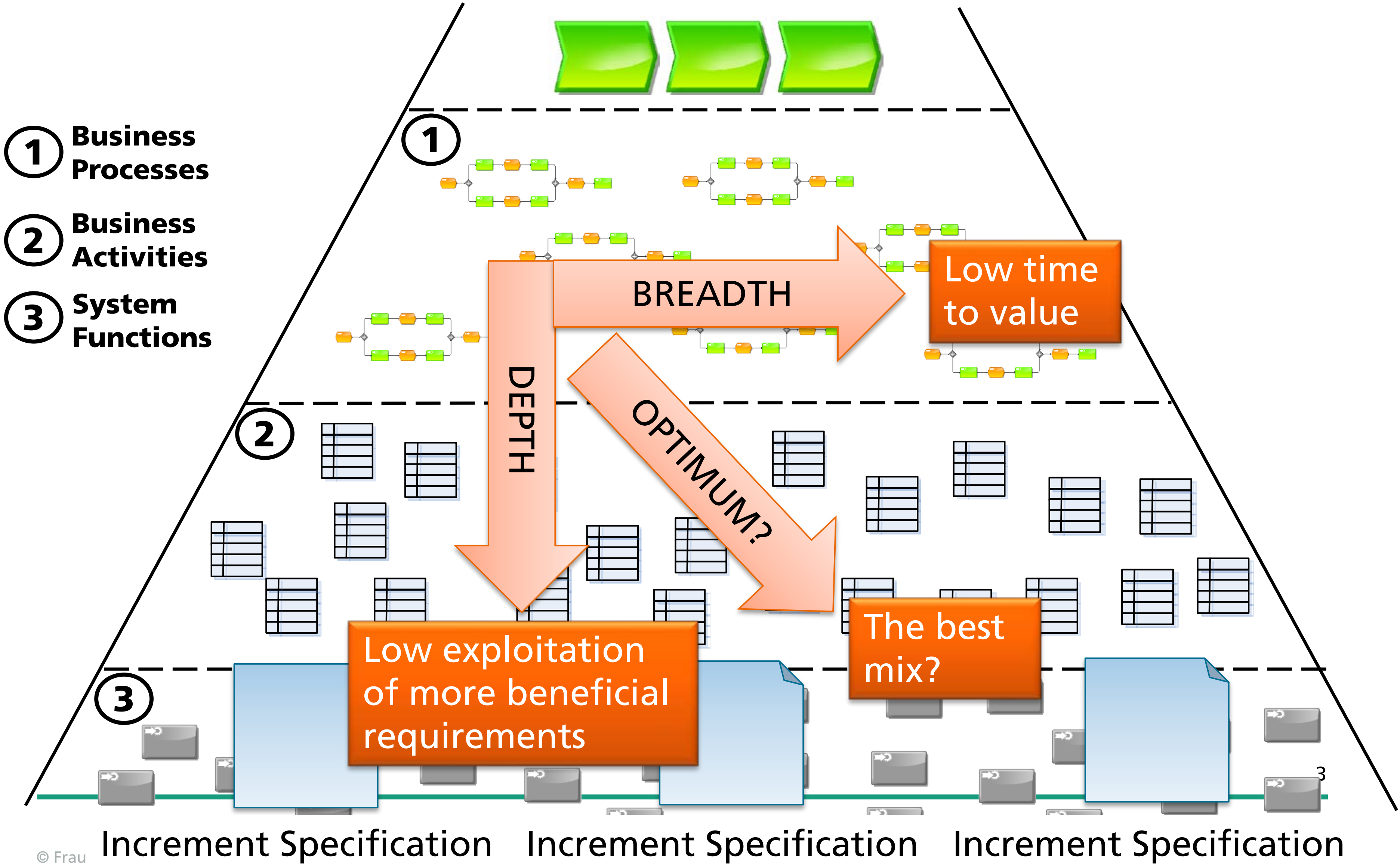
“Requirements engineers need to align requirements efforts to elicit and specify only the requirements that will deliver that value.”

Joy Beatty (Seilevel, USA)

Industry Keynote:
Stop Thinking About Requirements Quality, Focus on Value

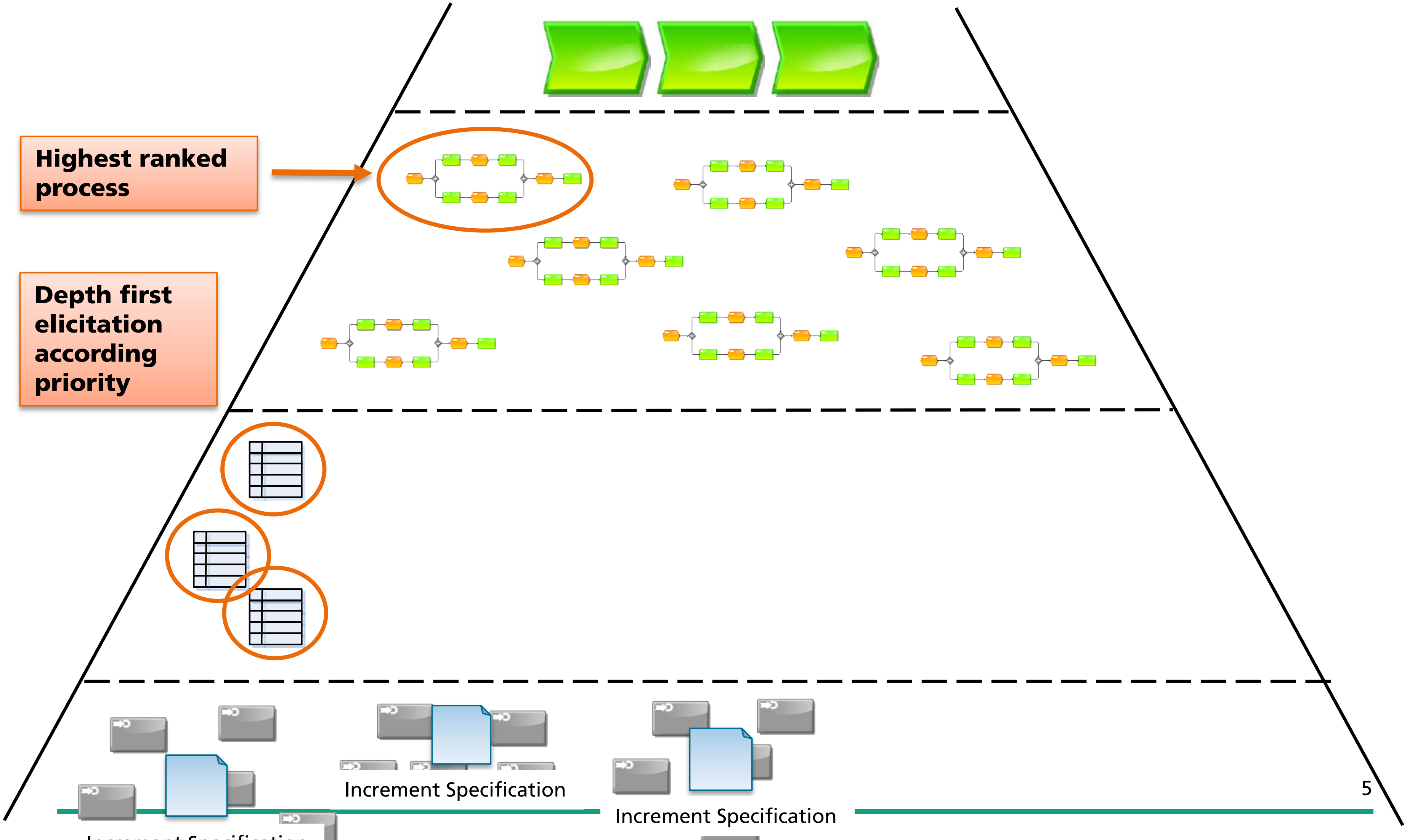
Fraunhofer IESE

Problems in Incremental Elicitation



IESE

Decision Heuristic: Depth First (based on priority)

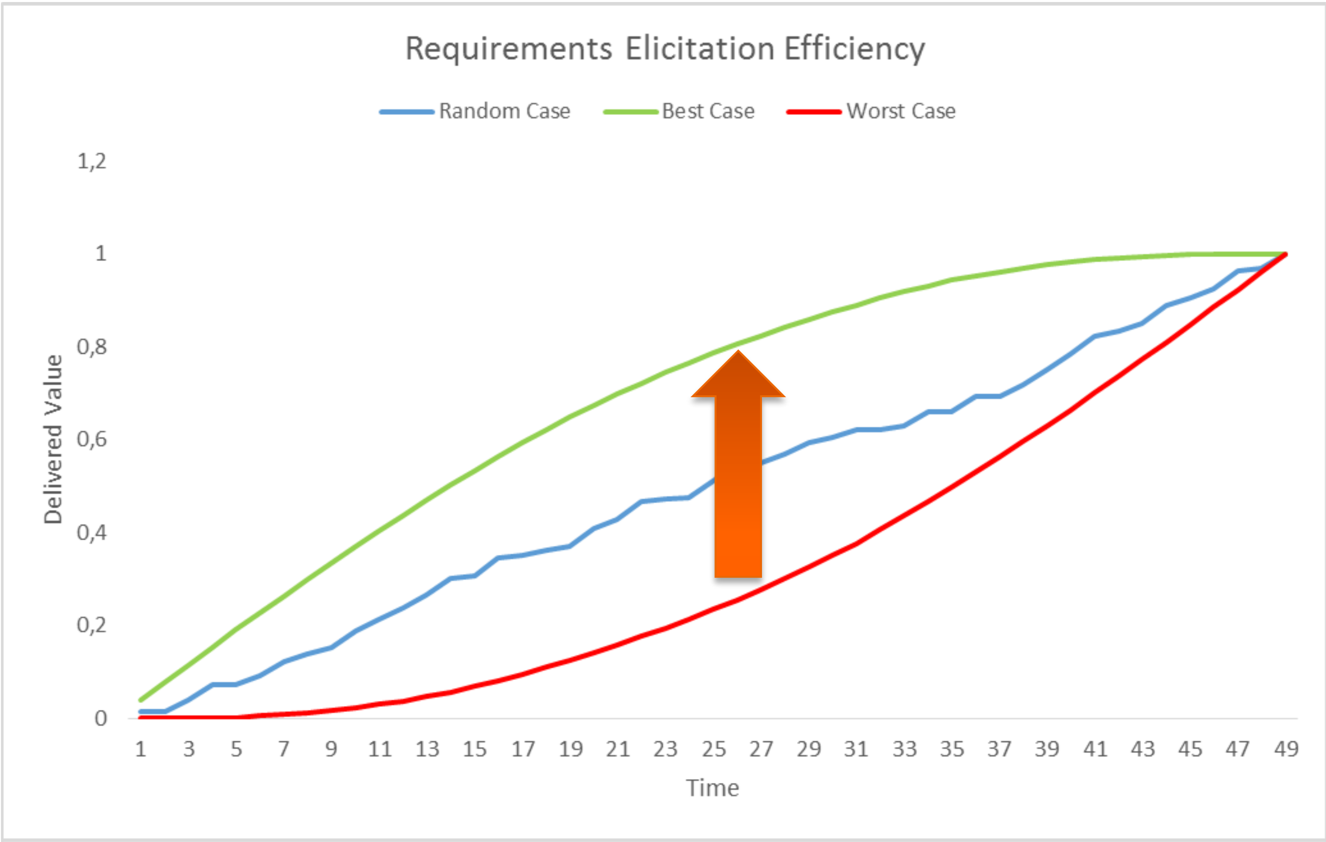


Fraunhofer IESE

Are there better Decision Heuristics?

Fraunhofer IESE

What is the Impact of different Decision Heuristics on Elicitation Efficiency?



Fraunhofer IESE

Research Questions

- **RQ$_{1:}$** *"Is there a difference between the elicitation efficiency when applying different decision heuristics in BPRE?"*

➢ **H$_{1,1}$** *There is a difference between elicitation efficiency in BPRE when applying different decision heuristics*.

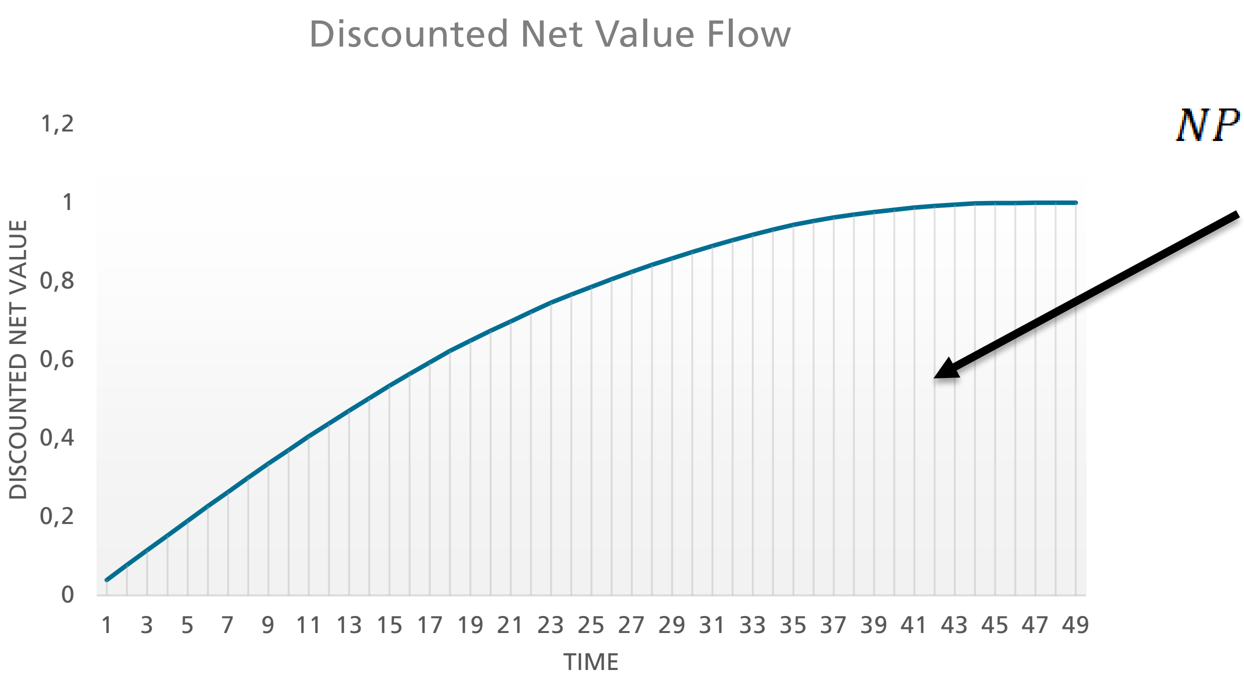
- **RQ$_{2:}$** *"Is there a difference between the elicitation efficiency when applying different decision heuristics in BPRE compared at different control points during project runtime?"*

➢ **H$_{2,1}$** *There is a difference between elicitation efficiency when applying different decision heuristics compared at different control points during project runtime*.

> Tool-based **simulation of** different **decision heuristics on** various business-process-based **requirements hierarchies** and analysis of their performance

Elicitation Efficiency Measure

- To express elicitation efficiency, the concept of **net present value (NPV)** is adapted that is able to express that
  - early value generation is more profitable than late value generation
  - lower elicitation effort is better than higher elicitation effort.



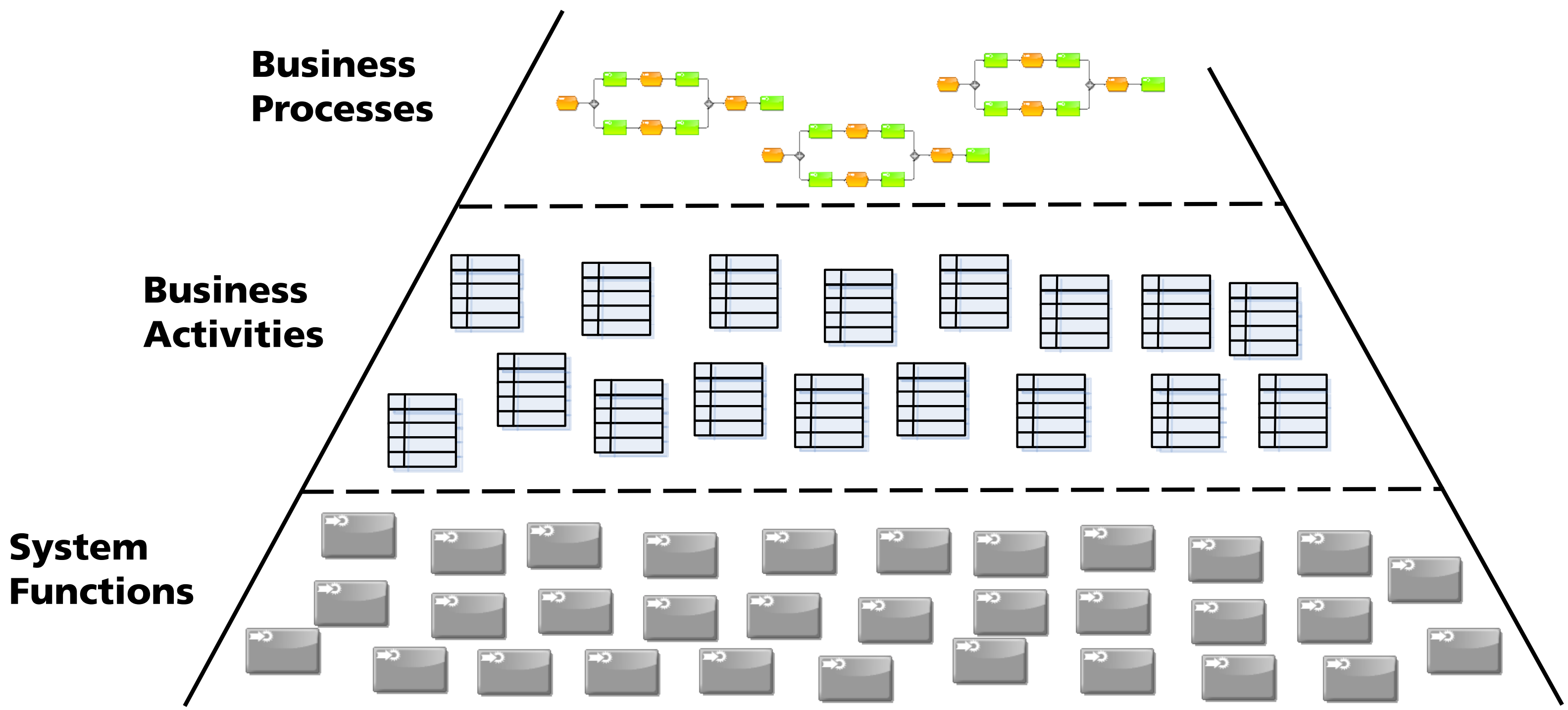
$$NPV(i,N) = \sum_{t=0}^{N} \frac{Z_t}{(1+i)^t}$$

<u>Assumption:</u>
Only System Functions generate value

Fraunhofer IESE

Requirements Model Parameters (1/3)

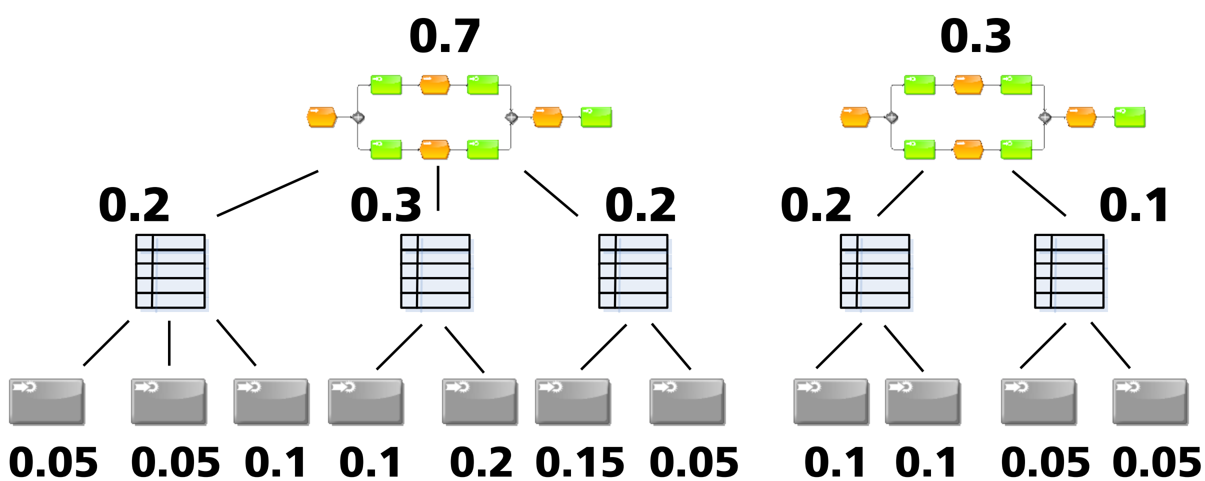
- Parameters based on past project experience & industry expert interviews
- **Three level requirements hierarchy** (BPs, BAs, SFs)
  - Requirements numbers normally distributed (e.g., mean BPs = 30)



Fraunhofer IESE

Requirements Model Parameters (2/3)

- **Priority values** randomly equally distributed and normalized
  - Equals Hierarchical Cumulative Voting (HCV) prioritization approach (Berander, 2006)



Requirements Model Parameters (3/3)

- **Elicitation Effort** for different requirements based on
  - **# of elicitation periods** (= hours needed for elicitation, e.g. interview time, workshop time)
  - **# of resource**s needed for elicitation (= persons needed per period)
  - BP and BA effort influenced by number of sub requirements (e.g., large process vs. small process, complex activity vs. simple activity)
  - Additionally, normal distributed for expressing variations

Fraunhofer IESE

Decision Heuristics

- Comparison of nine different decision heuristics (DH) based on priorities
  - SotA heuristics derived from literature (2 heuristics, e.g. DH1)
  - Adapted & newly created (7 heuristics, e.g. DH4)
    - 2 informed heuristics ("know" also elicitation effort)
- **Examples:**
  - **DH1. Highest Value (HV) First**
    - *At each decision point in the hierarchy, always refine the requirement with the highest priority next.*
  - **DH4: Remaining Value Global**
    - *Get the sum of the priorities of the most detailed requirements available in backlog; check if sum is greater than the highest priority of the requirement in the above hierarchy level; if yes, refine the low level requirement with the highest priority; if no, go to the higher hierarchy level and repeat procedure.*

Fraunhofer IESE

Experimental Procedure

- Two simulations
  - Simulation 1 for testing $\mathbf{H_{1,1}}$ (two-way repeated measures ANOVA)
  - Simulation 2 for testing $\mathbf{H_{2,1}}$ (one-way repeated measures ANOVA)
- Tool-based simulation of each decision heuristic on each of the generated requirements trees
- The tool automatically calculates the NPV at each control point (CP)

	Simulation 1 for testing $H_{1,1}$	Simulation 2 for testing $H_{2,1}$
Trees	25	75*5 (75 trees for 5 CPs each)
# Business Processes	813, Avg: 32.52	11472, Avg: 30.59
# Business Activities	16267, Avg: 650.68	229373, Avg: 611.66
# System Functions	81700, Avg: 3268	1150227, Avg: 3067,27
Runtime	01h:15m:40s, Avg: 03m:02s	22h:16m:16s, Avg: 03m:34s

Experiment Overall Results

- *Simulation 1:* **Mean elicitation efficiency differed statistically significantly** between all heuristics over all control points
  - $\mathbf{H_{1,1}}$ *"There is a difference between elicitation efficiency in BPRE when applying the decision heuristics"* **can be accepted**.

- *Simulation 2:* **Mean elicitation efficiency differed statistically significantly** between all heuristics **for CP1-CP5**
  - $\mathbf{H_{2,1}}$ *"There is a difference between elicitation efficiency when applying the decision heuristics compared at different points during project runtime"* **can be accepted**.

Fraunhofer IESE

Experimental Results: Simulation 1

- Post hoc tests using the Bonferroni correction revealed statistically significant differences between the heuristics

	DH1	DH2	DH3	DH4	DH5	DH6	DH7	DH8	DH9
DH2	41.16*	-							
DH3	165.68*	124.52*	-						
DH4	198.62*	157.46*	32.94*	-					
DH5	253.13*	211.96*	87.45*	54.51*	-				
DH6	124.81*	83.65*	-40.87*	-73.81*	-128.32*	-			
DH7	36.60*	-4.56	-129.08*	-162.02*	-216.52*	-88.21*	-		
DH8	192.70*	151.53*	27.02*	-5.92	-60.43*	67.89*	156.09*	-	
DH9	265.28*	224.12*	99.60*	66.66*	12.15	140.47*	228.67*	72.58*	-

**=statistically significant with p < 0.001*

Table 2. Pairwise Comparisons (i-j) of the Heuristics (Simulation 1)

- Decision Heuristics performance ranking:
  - **DH9 ~ DH5 > DH4 ~ DH8 > DH3 > DH6 > DH2 ~ DH7 > DH1**

Fraunhofer IESE

Experimental Results: Simulation 2

- Again, post hoc tests using the Bonferroni correction revealed statistically significant differences between the heuristics

  - **CP 1: DH5 > DH9 ~ DH4 > DH8 > DH3 ~ DH2 ~ DH6 > DH7 > DH1**
  - **CP 2: DH9 > DH5 > DH4 > DH8 > DH3 > DH6 > DH2 > DH7 > DH1**
  - **CP 3: DH9 > DH5 > DH4 ~ DH8 > DH3 > DH6 > DH7 ~ DH2 > DH1**
  - **CP 4: DH9 > DH5 > DH8 ~ DH4 > DH3 > DH6 > DH7 > DH1 ~ DH2**
  - **CP 5: DH9 > DH5 > DH8 ~ DH4 > DH3 > DH6 > DH7 > DH1 > DH2**

DH1-DH2 heuristics from literature

DH3-DH7 adapted & new heuristics

DH8-DH9 informed heuristics

DH1. Highest Value (HV) First
DH2: System Functions (SF) First
DH3: SF First-HV First
DH4: Remaining Value Global
DH5: Remaining Value Global All
DH6: Remaining Value
DH7: Remaining Value All
DH8 (informed heuristic): Value Cost Optimal
DH9 (informed heuristic): "Optimal" Solution

Interpretation (1/2)

- Decision Heuristics have the potential to make the requirements elicitation process more efficient
- Sophisticated heuristics perform best (DH5 “Remaining Value Globall All”)
- Intuitive decision heuristics seem to perform very low
- Informed heuristics which take into account elicitation effort
  - are even outperformed (esp. DH8) by heuristics only based on priorities (DH4 / DH5)
  - if no effort assessment is at hand, still good advice to base decisions on reasonably defined priorities
- Example with monetary numbers

DH5 NPV	DH9 NPV	DH4 NPV	DH8 NPV	DH3 NPV	DH6 NPV	DH2 NPV	DH7 NPV	DH1 NPV
58276$	54930$	53016$	42919$	36047$	27469$	20731$	19144$	10171$
(1447)	(1439)	(1432)	(1404)	(1387)	(1362)	(1339)	(1339)	(1314)

Table 3. Example: NPV for specific Setting of 14 Processes

Threats to Validity

- *Threats to Construct validity*
  - NPV: Assumption that value and effort can be directly compared
    - ➢ Usage of normalized values in the calculations, but hard to interpret
  - Model parameters of requirements trees
    - Realistic as possible, but mostly based on experience in past projects and expert interviews and not further validated
    - ➢ However, in further simulations with different tree sizes it seems that results can be reproduced
- *Threats to External validity*
  - Content of the requirements trees
    - Focus only on requirements which are directly derived from the processes
  - Reuse of requirements is not regarded (e.g., system functions that can be utilized in different business activities)
  - Assumption that value is directly generated after the elicitation of a system function is finished

Outlook

- Further simulations with different parameters done & planned
  - Different sizes for requirements trees (smaller, larger)
    - Almost similar (not statistically analyzed yet)
  - Different effort numbers (e.g., elicitation effort zero) → time to value
    - Almost similar (not statistically analyzed yet)
  - Different tree structure (e.g., only two levels)
  - Extension for different release time simulations
    - Normally distributed, packages, …
- Genetic algorithm for solving decision problem
  - Creates even better results than decision heuristics
- Integration in BPRE prioritization tool for use in industrial projects

Discussion

?