Kirkpatrick and beyond: A comprehensive methodology for influential training evaluations

Dr. Jan Ulrich Hense
Ludwig-Maximilians-Universität München, Department Psychology

Kirkpatrick's four levels of training evaluation

- Series of four articles 1959/1960

- Evaluating training programs 3rd ed. 2006

- Has been <u>the</u> dominant model for evaluation of training for decades.
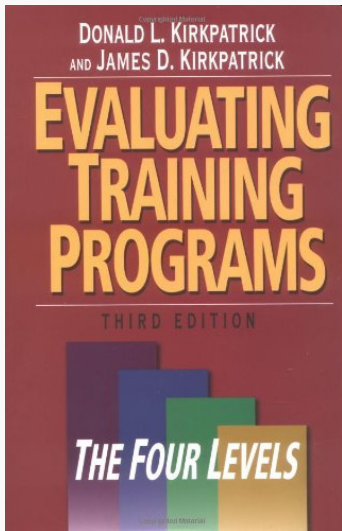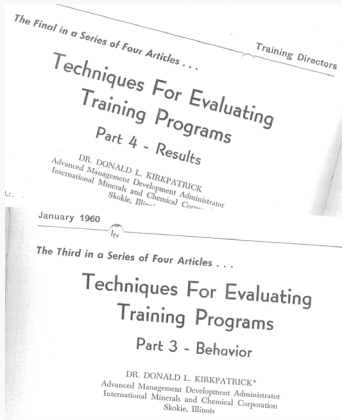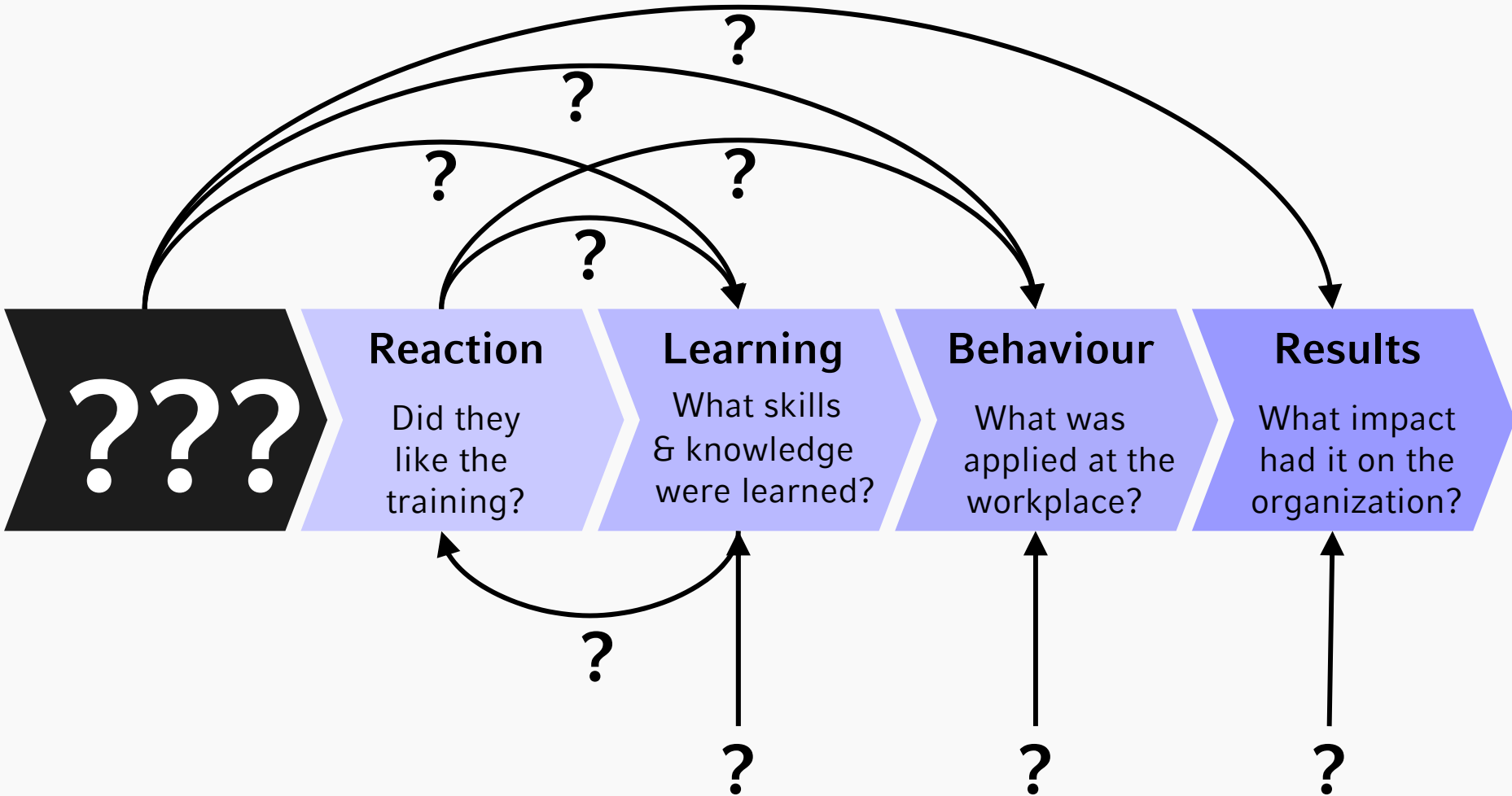
- Linearity
  - Theoretically not plausible (e.g. interaction between reaction and learning)
- Suggested causality ("chain of evidence")
  - Weak empirical evidence (e.g. Alliger et al., 1989; 1997)
  - Suggests reaction as a proxy for subsequent levels ("happy sheet evaluations")
- "Black-box" view of training
  - Narrow focus on outcomes (e.g. Holton, 1997)
  - Ignores training process, individual influences, and contextual influences on learning and transfer (e.g. Bates, 2004)
- "One size fits all" approach
  - Ignores variety of training specifics

- Limited use of four levels for influencing practice
  - Unable to provide formative information on training improvement
  - Risk of wrong summative conclusions

- Need for a more complete view of training reality
  - Inputs (e.g. participant characteristics)
  - Processes (e.g. learning methods)
  - Context (e.g. training and workplace conditions)

- Use of theory
  - Learning and instructional theory
  - Transfer theory
  - (cf. theory based evaluation, Weiss, Chen & Rossi, Bickman, Donaldson, Rogers ...)

Authentic context

Multiple contexts

Social context

Instructional context

Problem-based learning
Balance of construction
and instruction

Self-regulated learning

Learning with new media

Cooperative learning

(Reinmann & Mandl, 2006)

LMU

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Theory II:
Training transfer

Dr. Jan Hense

(Baldwin & Ford, 1988)
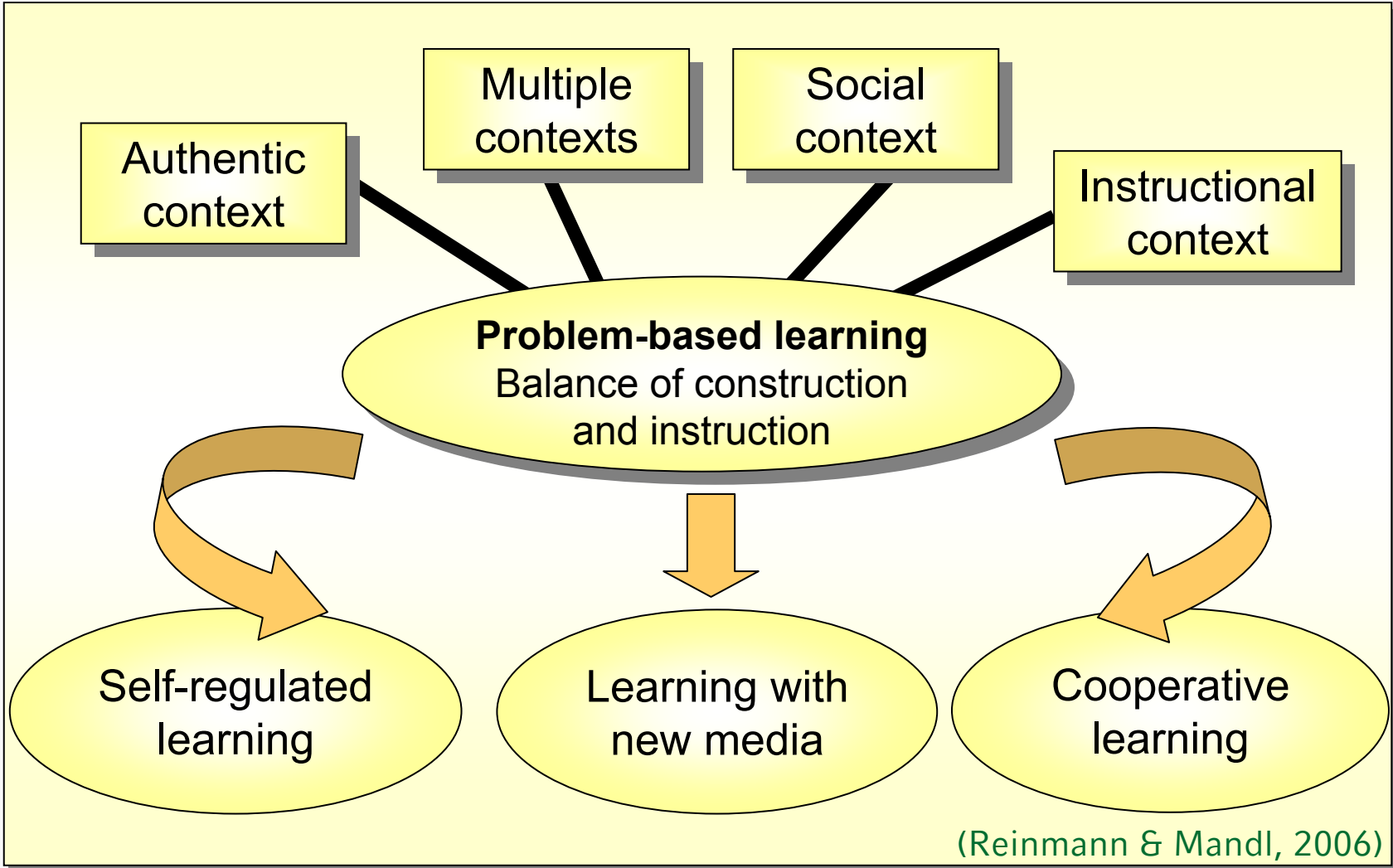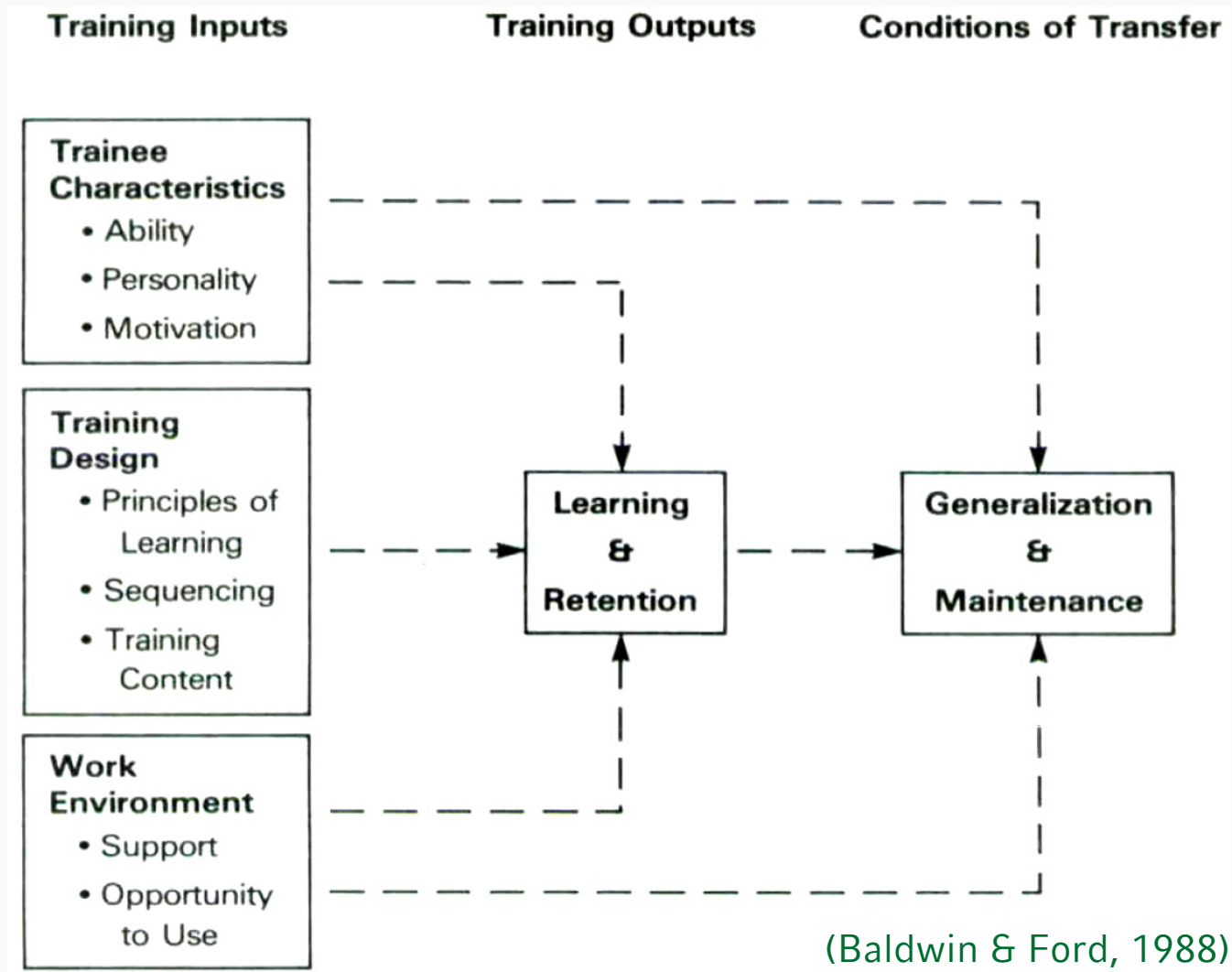
Evaluation of trainings for Six Sigma in a corporate setting

Six Sigma

- Systematic, data-driven approach to quality improvement
- Goal: error free processes

Methodology

- DMAIC cycle
- Related tool set

Roles

- Green & Black Belts (GB & BB)
- Master Black Belts (MBB)

Setting

- International corporation
- Specialized in industrial gases

DMAIC cycle

5. Control
How sustainable are the effects of improvements?

1. Define
What are the problems and goals?

2. Measure
What is the current state of things?

3. Analyze
What is wrong? What's causing the problem?

4. Improve
What can be done to improve the problem?

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Reaction → Learning → Behaviour → Results

Inputs	Training process	Outcomes	Transfer	Results
Training concept Goals Contents Design Methods	**Didactics** Problem-based learning Cooperative learning Self-directed learning	**Reaction** Satisfaction Usefulness	**Usability** of contents	Added value
Participants Learning preferences Motivation Previous knowledge Previous experiences Preparation	**Roles** Learner activities Trainer activities	**Immediate effects** Skills & knowledge Attitudes Self-efficacy	**Application** of skills and knowledge	Cost reduction
	Realization Learning sequences Media and materials	**Long-term effects** Project finalization Career options	**Support** by coach by manager	Cost-benefit ratio

Context

Workplace Need for training Line manager support	Training context Facilities Resources	Workplace Learning conditions Transfer conditions	Organi-zation

Study I: **Analysis of training concept**

- Sample: Manuals for trainers and participants
- Method: Structured document analysis & Interviews

Study II: **Training process observation**

- Sample: ca. 80 training hours (BB-training + GB-training)
- Method: Structured observation by two independent observers (event sampling with category scheme)

Study III: **Pre- & posttest surveys**

- Sample: 40 participants of ongoing GB- and BB-courses
- Method: Written surveys before and after the training

Study IV: **Transfer study**

- Sample: 34 MBBs & 77 former BB-participants
- Method: Online transfer surveys
- Additional interviews with MBBs and former BB-participants
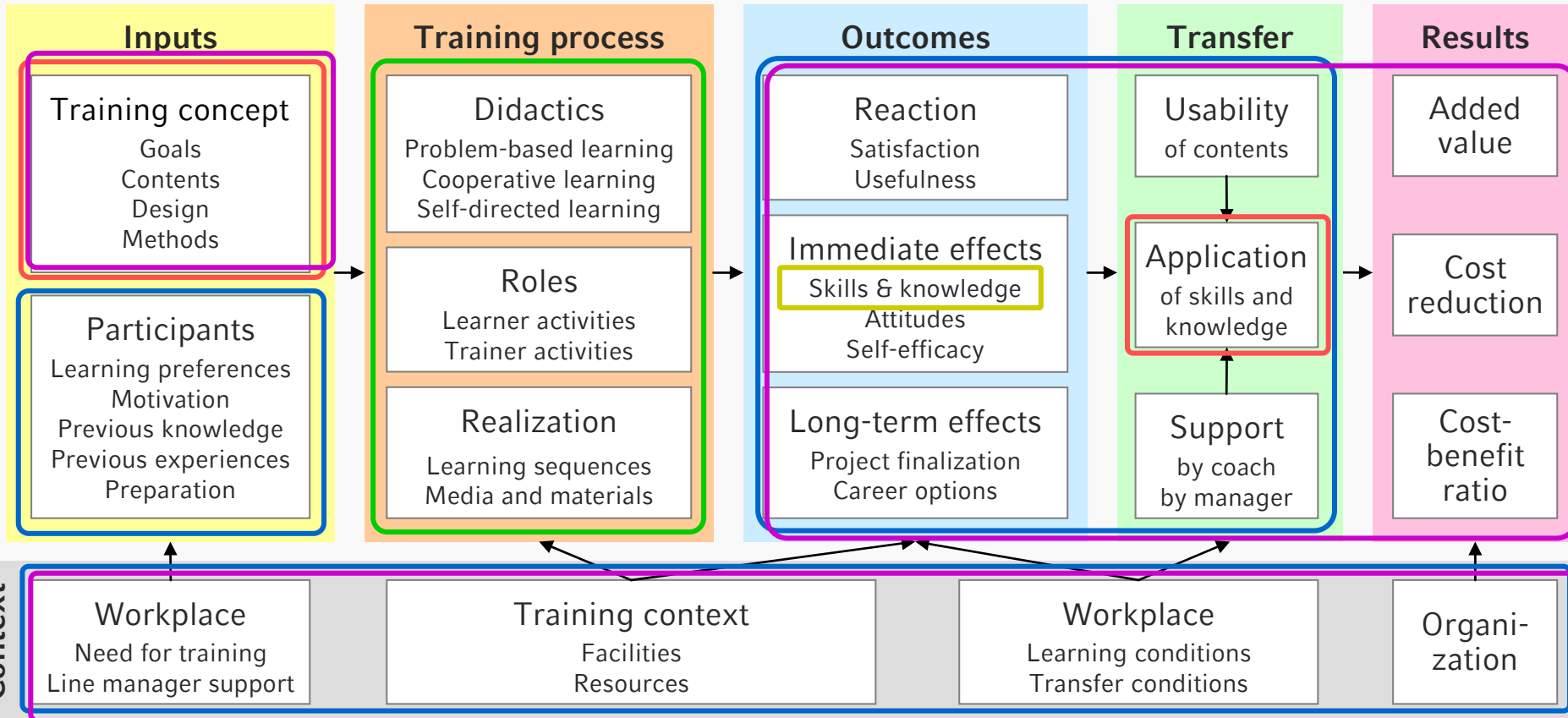- Additional analysis of finished Six Sigma projects

Document analysis | Observation | Test | Interviews | Surveys

Inputs

Training concept
Goals
Contents
Design
Methods

Participants
Learning preferences
Motivation
Previous knowledge
Previous experiences
Preparation

Training process

Didactics
Problem-based learning
Cooperative learning
Self-directed learning

Roles
Learner activities
Trainer activities

Realization
Learning sequences
Media and materials

Outcomes

Reaction
Satisfaction
Usefulness

Immediate effects
Skills & knowledge
Attitudes
Self-efficacy

Long-term effects
Project finalization
Career options

Transfer

Usability
of contents

Application
of skills and
knowledge

Support
by coach
by manager

Results

Added value

Cost reduction

Cost-benefit ratio

Context

Workplace
Need for training
Line manager support

Training context
Facilities
Resources

Workplace
Learning conditions
Transfer conditions

Organization

Six Sigma Training Evaluation
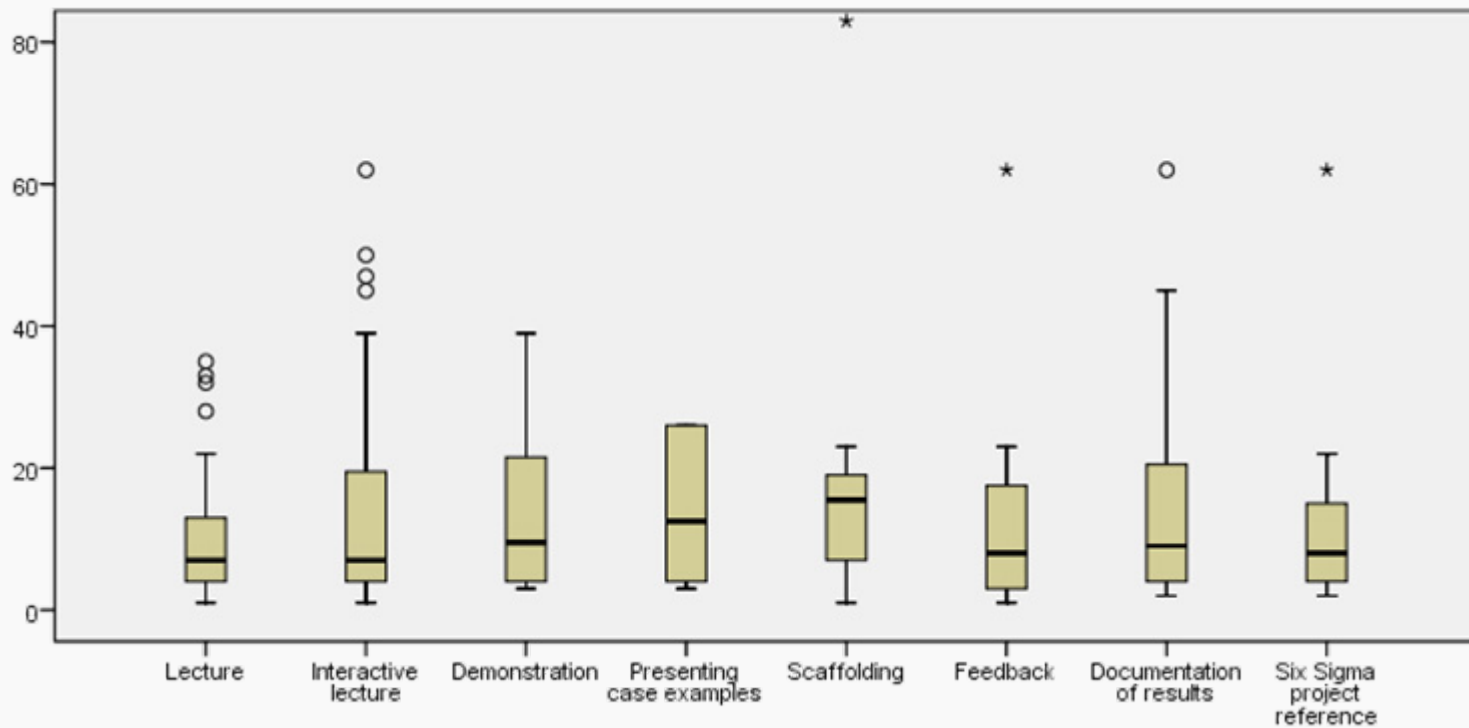## Category scheme for training observation

1. Categories for trainer activities
2. Categories for social learning forms
3. Categories for learner activities
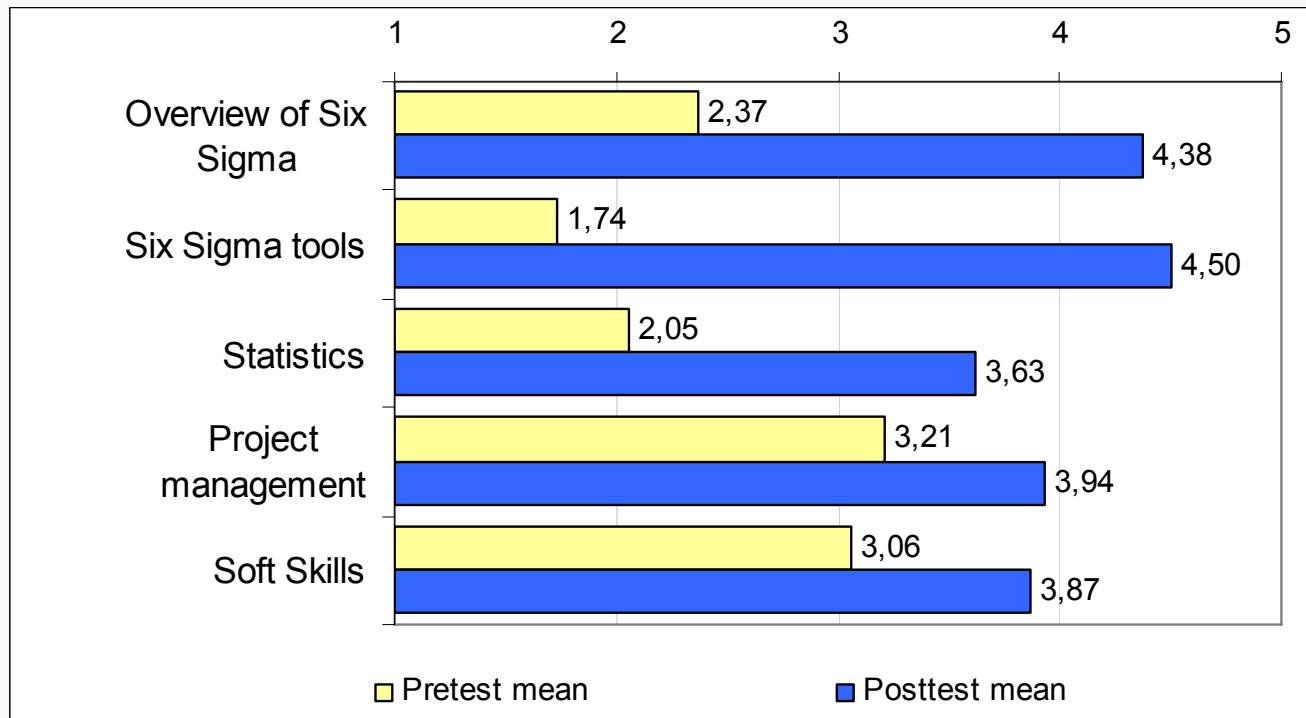4. Categories for media use

1. Categories for trainer activities

Category	Operational definition	Rationale
Lecture	The trainer presents contents in a *unidirectional* way without encouraging active participation of learners.	Lectures are probably the most common form of trainer activity. They are specifically suited for introducing new topics. Due to the nature of the human attention span, they should be used sparsely and in time chunks not longer than 10 to 15 minutes. Since learners have a passive role during lectures, their interest and motivation are particularly important for effective learning to take place.
Interactive lecture	The trainer presents contents in a *dialogical* way which encourages learners to participate actively, e.g. by asking questions.	Interactive lectures prolong learners' attention span and potentially deepen their cognitive involvement by encouraging their active participation. However, since usually only a small number of learners actually engage in answering questions etc., care has to be taken if indeed all learners are activated by interactive lectures. Accordingly, interactive lectures, too, should usually not be longer than 15 minutes.
Demonstration	The trainer explains how to use or apply specific materials, tools	On the surface, demonstrations resemble lectures, as they are mainly unidirectional. However, since they work with authentic materials

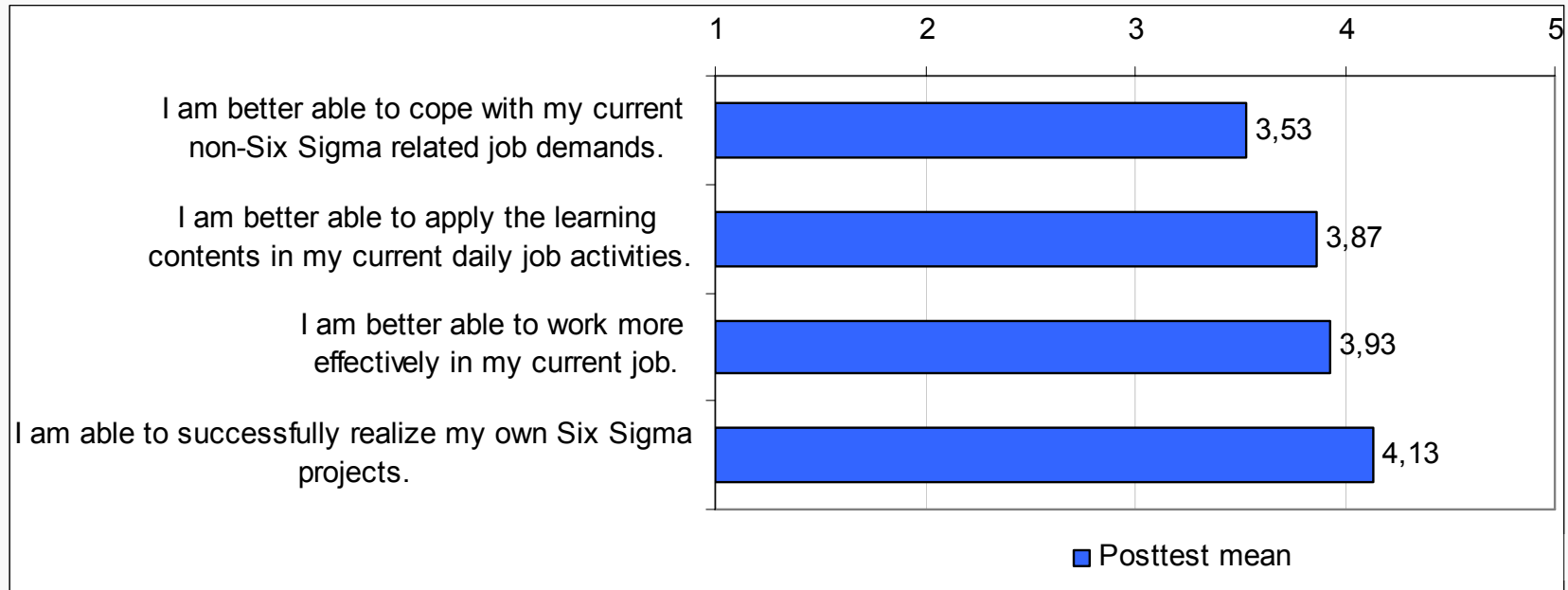| Start	Modul	Modul Nr.	Trainer (Initialen)	Traineraktivität								Soziale Form		Lerneraktivität								Medien des T...		
				Lehrvortrag	Lehrgespräch	Demonstration	Beratung/Unterstützung	Präsentation Fallbeispiel	Ergebnissicherung	SixSigma Projektbezug	Rückmeldungen	Einzelarbeit	Gruppenarbeit	Präsentieren	Übung	arbeiten am Fallbeispiel	Diskussion	stellen Fragen	Planspiel/ Rollenspiel	Six Sigma Projektbezug	Energizer/Warm-up	Powerpoint/Beamer	FlipChart	Pinnwand
08:30	Intro to day 3; Agenda		HR		X																			
08:36	Recap Day 2 exercise												x		x									
08:52	report out recap exercise						x				x													
09:11	exercise: Leader vs. Facilitator				X																		x	
09:19	Coffe Break 10min																							
09:31	Graphical analysis and baic statistics	2212	XX		X													x					x	
09:52	Minitab installation, Notebooks einrichten						x					x												
10:08	Minitab intro				x	x	t	t				x						x				x	x	
11:05	Coffee Break																							
11:25	planning data collection				x																		x	
11:27	cont. Minitab intro				x	x		t				x						x				x	x	

	Lecture	Inter-active lecture	Demon-stration	Presen-ting case exam-ples	Scaffol-ding	Feed-back	Docu-men-tation of results	Six Sigma project reference
Number of sequences	35	51	16	6	18	15	12	9
Average min. per sequence	10:34	13:44	13:04	24:50	29:13	13:08	16:30	15:13

Selected results:
Study III: Pre- & posttest surveys
(BB-training participants)

Dr. Jan Hense

Skills & Knowledge (pre & post test)

Bar chart showing pretest and posttest means on a scale from 1 to 5:

Category	Pretest mean	Posttest mean
Overview of Six Sigma	2,37	4,38
Six Sigma tools	1,74	4,50
Statistics	2,05	3,63
Project management	3,21	3,94
Soft Skills	3,06	3,87

☐ Pretest mean ■ Posttest mean

Selected results:
Study III: Pre- & posttest surveys
(BB-training participants)

Dr. Jan Hense

Self-efficacy (post test)

	Posttest mean
I am better able to cope with my current non-Six Sigma related job demands.	3,53
I am better able to apply the learning contents in my current daily job activities.	3,87
I am better able to work more effectively in my current job.	3,93
I am able to successfully realize my own Six Sigma projects.	4,13

■ Posttest mean

Selected results:
Study III: Pre- & posttest surveys
(BB-training participants)

Dr. Jan Hense

Project finalization (post test)

	0%	20%	40%	60%	80%	100%
Define	7,00			8,00		
Measure	1,00	10,00			4,00	
Analyze	5,00		10,00			
Improve	1,00	14,00				
Control	15,00					

□ finished   ■ ongoing   ■ not started yet
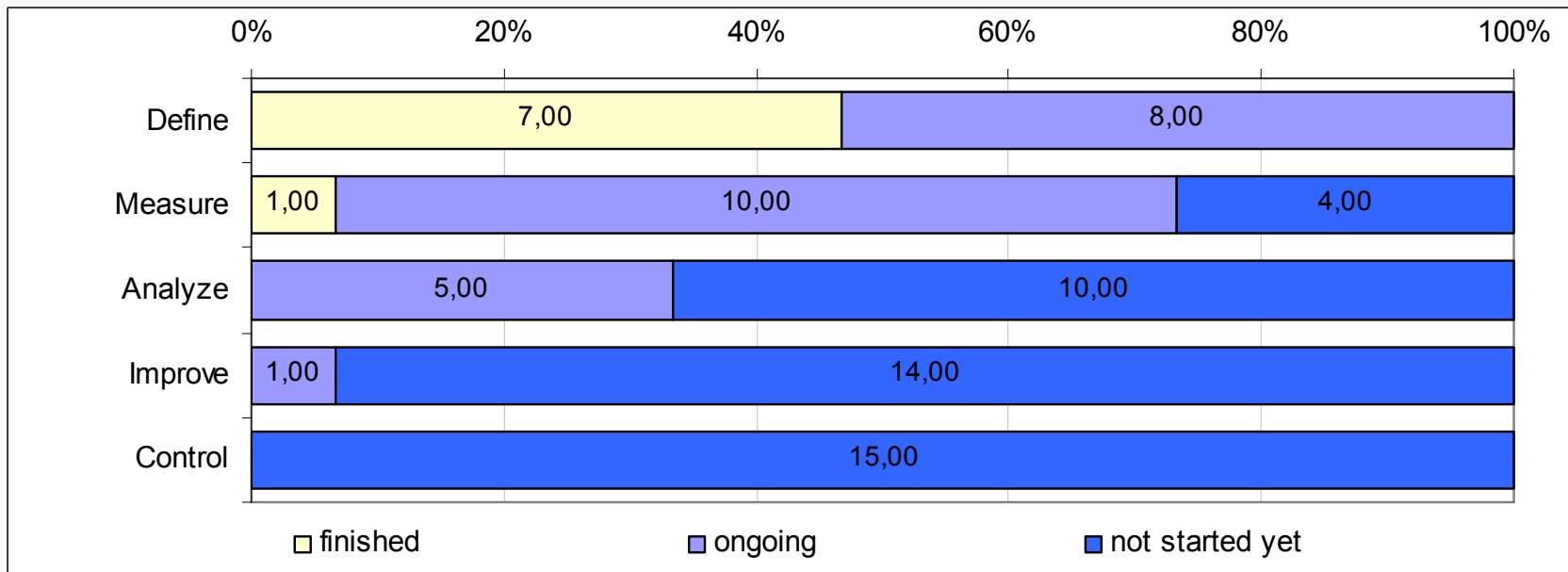
(number of participants)

Transfer conditions
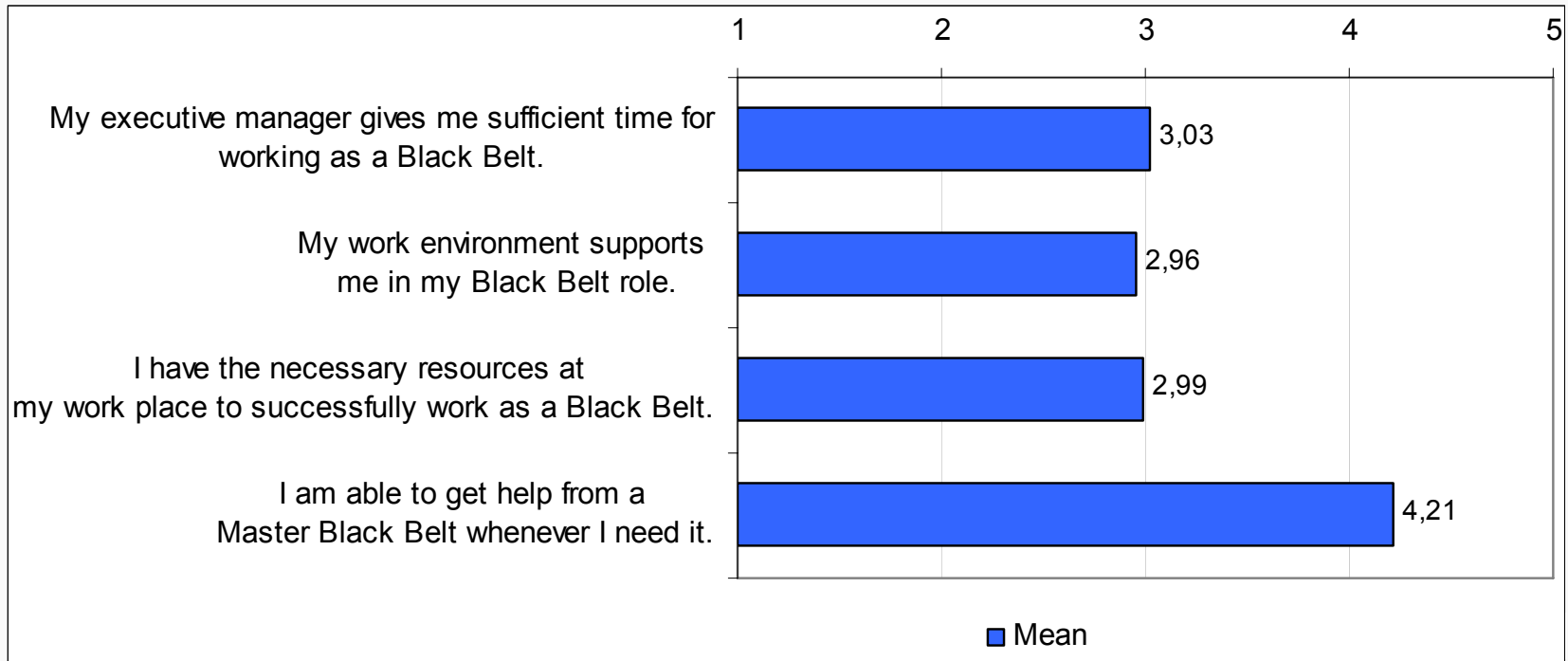
Selected recommendations to training management:

Input

- Improve preparation of participants
  (e.g. obligation to define own project with MBB)
- Slightly reduce training contents

Process

- Reduce number and length of (interactive) lectures
- Refer to learners' own projects more regularly
- Synchronize training weeks with project progress

Context

- Strengthen transfer conditions
  (e.g. line manager support, post-training workshops)
- Process control and support of participants' projects
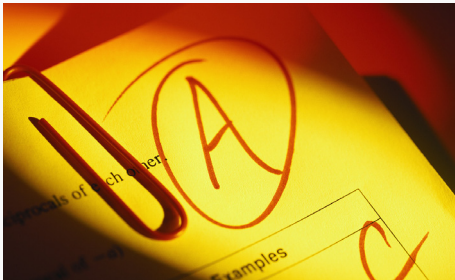
Merits of the evaluation approach

- The multi-method approach, comprising
  - input,
  - process,
  - outcome, and
  - transfer analyses,

  generated multiple formative insights into possible improvements of the trainings' concept and implementation.

- This would not have been possible with an entirely outcome-oriented evaluation.

On a **general evaluation theory and policy level**, the study can serve as an example of the limitations of pure impact evaluations for influencing practice.

In many practical contexts they will not provide the information necessary to guide **improvements** of a program.

Inclusion of **theoretical considerations** on input, context and process influences is often indispensable for evaluation to sustainably influence practices and policies.

Thank you for your attention!

jan.hense@psy.lmu.de

Alliger, G. M. & Janak, E. A. (1989). Kirkpatrick's level of training criteria. Thirty years later. *Personnel Psychology*, *42*, 331-342.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H. & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, *50*, 341-358.

Baldwin, T. & Ford, K. (1988). Transfer of training: A review and directions for future research. *Personal Psychology*, *41* (1), 63-105.

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, *27*, 341-347.

Holton, E.F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, *7*, 5-21.

Kirkpatrick, D. (1959a). Techniques for evaluating training programs. Part 1 - Reaction. *Journal of the American Society of Training Directors, 13 (11),* 3-9.

Kirkpatrick, D. (1959b). Techniques for evaluating training programs. Part 2 - Learning. *Journal of the American Society of Training Directors, 13 (12),* 21-26.

Kirkpatrick, D. (1960a). Techniques for evaluating training programs. Part 3 - Behavior. *Journal of the American Society of Training Directors, 14 (1),* 13-18.

Kirkpatrick, D. (1960b). Techniques for evaluating training programs. Part 4 - Results. *Journal of the American Society of Training Directors, 14 (2),* 28-32.

Kirkpatrick, D. L. & Kirkpatrick, J. D. (2006). *Evaluating training programs. The four levels (3rd ed.)*. San Francisco: Berret-Koehler.

Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten [Teaching and design of learning environments]. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie. Ein Lehrbuch [Educational psychology. A textbook]*. 5. ed. (pp. 613–658). Weinheim, Germany: Beltz.