Pattern-driven Process Design

Michael Zapf

Working Paper 7/2003

Working Papers in Information Systems 1

University of Mannheim
Information Systems 1
D-68131 Mannheim/Germany
Phone ++49 621 1811691, Fax ++49 621 1811692
E-Mail: wifo1@uni-mannheim.de
Internet: http://www.bwl.uni-mannheim.de/wifo1

Pattern-driven Process Design

Michael Zapf[*)], University of Mannheim

Abstract. With the pattern-driven process design approach we provide an instrument for the design of business processes which is based on objective performance criteria. Since enterprises are more and more confronted with an extreme dynamic environment a technique for measuring the flexibility of business processes is incorporated which enables process designers to chose the most flexible process for their purpose. The pattern-driven process design has been applied within the communication center domain. Two types of domain specific process patterns have been identified regarding the qualification-mixture and communication-mixture within the organization. These real life patterns have been evaluated on the base of empirical data from four different lines of business: car rental, bank, book trade and energy industry.

Keywords. Business process design, process mining, process flexibility, process patterns, simulation, communication center

---

[*)] Michael Zapf, Information Systems 1, University of Mannheim, Schloss,
D- 68131 Mannheim/Germany, Phone: ++49 621-181-1693, Fax: ++49 621-181-1692,
E-Mail: zapf@bwl.uni-mannheim.de

1 Introduction

For business process management many tools, modeling techniques and procedure models have been developed during the last years (e.g. Van der Aalst et al., 2000, and Becker et al., 2001). But still little knowledge exists how to design processes from a normative prospective. One approach for this design question is the development of reference models for specific lines of business which can be used as starting point for individual process models (e.g. Schütte, 1998; Becker at al., 1996 and Scheer, 1998). Within this approach empirical cases are identified but no evidence is given for the superiority of these reference processes. The process designer is not able to estimate in advance, whether the utilization of a reference process leads to the desired positive results.

With the pattern-driven process design approach we provide an instrument for the design of business processes which is based on objective performance criteria. According to the increasing importance of service speed the process efficiency plays an important role in our approach. Since enterprises are confronted with an extreme dynamic environment and increasing individualization of customer requirements the flexibility of processes gets more and more important. We incorporated a technique for measuring the flexibility of business processes and enable process designers to chose the most flexible process for their purpose.

The pattern-driven process design has been applied within an experimental study in the communication center domain. Two types of domain specific process patterns have been identified regarding the qualification-mixture and communication-mixture within the organization. The qualification-mixture defines the combination of agents with different qualification levels whereas the communication-mixture comprises alternative strategies for organizing communication channels. The real life patterns have been evaluated on the base of empirical data from four different lines of business: car rental, bank, book trade and energy industry.

Section 2 gives an overview of related contributions from the literature. In the next section the pattern-driven process design is presented in detail with a procedure model and a technique for measuring process flexibility. Section 4 describes the practical application in the communication center domain. The relevant process patterns and performance measures are identified and the research hypothesis are formulated. After that the empirical data and the experimental design is presented. The results are outlined and discussed. In the closing section the results are summarized, limitations are discussed and future research directions are derived.

2 Related contributions

The concept of pattern-driven process designs traces back to Hammer et al. (1990), who presented some general business process patterns and design guidelines but did not give evidence for their recommendations. Buzacott (1996), Dewan et al. (1997) and Seidmann et al. (1997) evaluate some of these patterns with queuing theory respectively linear programming. Because of the limitation of these methods concerning the modeling complexity, only simple patterns have been analyzed under strong restrictions. The pattern-driven process design uses simulation to overcome these restrictions and allows therefore the evaluation of process patterns close to reality.

Another characteristic of pattern-driven process design is the evaluation of process flexibility. Hereby it is examined whether the process efficiency remains sufficient even if the environmental parameters change. For measuring flexibility we develop a new approach which is based on the ideas of Jacob (1974), Hanssmann (1987) and Kühn et al. (1990).

The coordination theory from Thompson (1967) is another source for the deduction of design guidelines for business processes. Thompson differentiates between (1) reciprocal, (2) sequential and (3) pooled interdependencies within an organization and recommends to build organizational units according to interdependencies (1) and (2). Some other approaches like Kilman (1987), Crowston (1997) and Malone et al (1999) use coordination theory as basis for deriving business process typologies and design methods. In this paper coordination theory is used to derive and substantiate the research hypothesis for the practical example from the communication center domain.

For the first step we applied the pattern-driven process design approach for relevant design patterns in the communication center domain. Van der Aalst et al. (2003) established a repository of general workflow patterns and with that created a more general basis for analysis of design patterns. So far these patterns have used for structural comparison and improvement of workflow management systems but in future they can be used as building blocks for domain specific patterns and as a starting basis for pattern-driven process design.

Another way to identify relevant process patterns is given by workflow mining (e.g. van der Aalst et al., 2002). In this approach the pattern structure is automatically derived e.g. from workflow management or enterprise resource planning systems. On this basis the pattern-driven process design can be utilized to evaluate the current pattern, compare it with another generic pattern and create a feedback loop for automatically improvement of the current process.

3 Pattern-driven process design

3.1 Overview

Within the pattern-driven process design approach, one base pattern is selected from a group of available patterns concerning a given set of performance measures and current process information (Figure 1). The selected base pattern serves as starting point for the future process design.

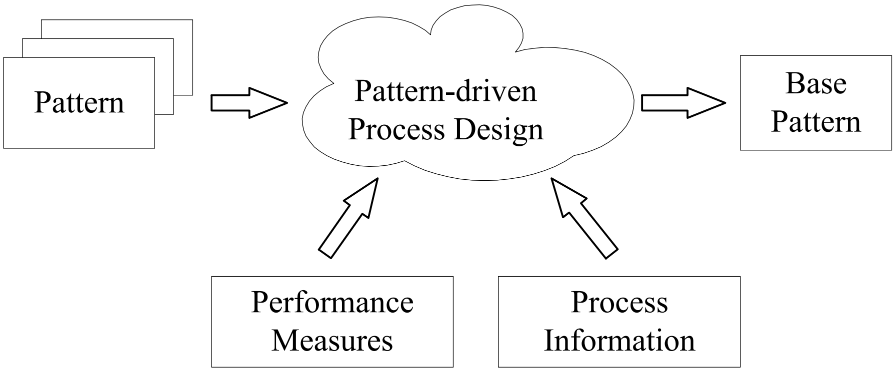

Fig. 1. Pattern-driven process design

As prerequisite for this approach all patterns have to be applicable for the intended process objective and their performance has to be measurable with the same performance measures. A process pattern for fast handling of technical support requests may not be comparable to a pattern for processing secure purchasing transactions over the internet. In the first case the processing time is the primary performance indicator whereas the security level may be an indicator for the second pattern.

Although the same performance measures have to applicable to the different patterns they normally attain different overall process performance in different business environments. With the help of process mining techniques information about the current business environment (process information) can be detected and used as basis for the appropriate design decisions.

3.2 Performance measurement

In every business domain specific performance measures are used. For example may be the processing time an appropriate measure for the production process of automobile industry. The measures are calculated based on a predefined set of input values (environment), like the number of cars produced per hour and the average processing time per single process activity. Normally the attained performance changes if the environment changes and a process may be efficient in one environment but not efficient in an other environment. So it is not a good idea to take only one combination of input values into consideration when making a design decision. Therefore a new approach for calculating the flexibility of a process has been developed in order to aggregate the performance measures under different environmental situations (Zapf, 2001). The process flexibility can be calculated with domain-specific performance measures and helps to overcome the shortcomings of traditional performance analysis approaches.

3.2.1 Partial flexibility values

The process flexibility is a comparative measure which will be calculated by comparing the performance of different designs. The building blocks are called partial flexibility values f_{ij} which are calculated for every design A_i and every analyzed environmental situation U_j concerning a given performance measure (see Zapf, 2001 for a detailed derivation of the flexibility measure):

$$f_{ij} = 1 - \frac{x_{ij} - x_{jmin}}{max(max_l(x_{lj}), s_j) - x_{jmin}} \tag{1}$$

The values are standardized with the minimum performance value $x_{j\,min}$ per environmental situation U_j. In order to avoid distortions if all designs have small performance values, predefined threshold values s_j are used for every environmental situation U_j. Keep in mind that the definition bases on the assumption that lower performance values are preferable to higher values which is applicable e.g. for processing times. With little transformation the flexibility values can also be defined for performance measures with an other characteristic.

	U_1	U_2	U_3
$x_{j\,min}$	3.0	3.0	1.4
s_j	4.0	4.0	2.0
$x_{1,j}$	4.0	4.8	1.7
$x_{2,j}$	4.0	6.0	1.5
$f_{1,j}$	$1 - \frac{4.0 - 3.0}{4.0 - 3.0} = 0.0$	$1 - \frac{4.8 - 3.0}{6.0 - 3.0} = 0.4$	$1 - \frac{1.7 - 1.4}{2.0 - 1.4} = 0.5$
$f_{2,j}$	$1 - \frac{4.0 - 3.0}{4.0 - 3.0} = 0.0$	$1 - \frac{6.0 - 3.0}{6.0 - 3.0} = 0.0$	$1 - \frac{1.5 - 1.4}{2.0 - 1.4} = 0.83$

Table 1. Partial flexibility values

Table 1 gives an numerical example of partial flexibility values concerning the average turnaround time of customer requests for two call center designs in three different environmental situations. This example shows no difference of the designs in U_1, an advantage of A_1 in situation U_2 and weaknesses of A_1 in U_3.

3.2.2 Aggregate flexibility values

For taking multiple environmental situations into account we build aggregate flexibility values which can be calculated in different ways depending on the degree of information about the environmental situations. In this paper we focus on the situation of uncertainty and assume that every environmental situation occurs with the same probability. This assumption leads to an aggregation with equal weights. If more information about the occurrence of single environmental situations is available, various weights can be used (see Zapf, 2001, for more details). The aggregate flexibility value for design A_i in n environmental situations is calculated as:

$$F_i = 1 - \frac{\sum_{j=1}^{n} (x_{ij} - x_{jmin})}{\sum_{j=1}^{n} (max(max_l(x_{lj}), s_j) - x_{jmin})} \qquad \textbf{(2)}$$

With the numerical values of Table 1 we get the following aggregate flexibility values F_1 and F_2 for design A_1 and A_2:

$$F_1 = 1 - \frac{(4-3)+(4.8-3)+(1.71-1.4)}{(4-3)+(6-3)+(2-1.4)} = 0.32 \qquad \textbf{(3)}$$

$$F_2 = 1 - \frac{(4-3)+(6-3)+(1.5-1.4)}{(4-3)+(6-3)+(2-1.4)} = 0.11$$

The partial and aggregate flexibility values can be presented in a common flexibility matrix (Table 2).

	U_1	U_2	U_3	F_i
$x_{j\,min}$	3.0	3.0	1.4	
s_j	4.0	4.0	2.0	
$x_{1,j}$	4.0	4.8	1.7	
$x_{2,j}$	4.0	6.0	1.5	
$f_{1,j}$	0.0	0.4	0.5	0.32
$f_{2,j}$	0.0	0.0	0.83	0.11

Table 2. Flexibility matrix

4 Practical application in the communication center domain

4.1 Design dimensions and process patterns

Within a communication center organization two basic activities have to be performed in order to handle incoming customer requests successfully:

- classify incoming request: accept request and if necessary forward it to a suitable qualified employee and
- handle request.

The classifying activity is used to partition the overall call volume and provide separate process versions and resources for each partition (e.g., Hammer et al., 1993 and Zapf et al., 2000). In practice the call volume is often partitioned according to specialization and communication reasons. Therefore the two main designs dimensions qualification-mixture and communication-mixture have been identified (Zapf, 2001)[1].

Incoming customer requests have different levels of difficulty. The design patterns for the qualification-mixture dimension distinguish between standard and special requests (Zapf et al., 2000 and Zapf et al., 2001). Standard requests deal for example with the processing of simple transactions, the modification of customer data or general enterprise or product information. They are normally handled by employees (agents) with basic knowledge. These agents will be called generalists. Some other requests refer to difficult technical problems, extensive consultations or complaints. They are called special requests and can only be handled by specialists who have specific, in-depth knowledge or special skills. The main design question for the qualification-mixture is: What is the appropriate mixture of generalists and specialists for handling a given volume of standard and special requests?

The deployment of different communication channels like phone, e-mail, fax or chat lead to another type of process partitioning according to the channel which is used by the customer. In this area we distinguish between synchronous and asynchronous channels which can both be used for standard and special requests. Synchronous communication takes place if all parties are communicating with each other at the same time (e.g., phone or chat). E-mail and fax are examples for asynchronous communication channels, where normally longer time intervals pass by between the single communication steps. The communication partners do not need to get in contact at the same time. The assignment of requests from a particular channel to the suitable agent is the main design question in the communication-mixture dimension.

4.1.1 Qualification-mixture patterns

The qualification-mixture will be defined by the assignment of process activity to agent group for each request type. The assignments are shown in Table 3.

group	**generalist group**		**specialist group**	
activity	**classify**	**handle**	**classify**	**handle**
2-level	all standard	all standard		
	all special			all special
back-office	syn. standard	syn. standard	asyn. standard	asyn. standard
	syn. special		asyn. special	all special
1-level	all standard	all standard	all standard	all standard
	all special		all special	all special

Table 3. Activities and agent groups (legend: syn. = synchronous, asyn. = asynchronous)

Within the classical 2-level pattern the internal communication structure is divided up into a first and a second level (Figure 2)[2]. Generalists accept all incoming requests. Standard requests are directly handled by the accepting generalist in the first level. Special requests are forwarded to an specialist agent in the second level. In the case of complex tasks this pattern will be expanded with further levels in practice.

---

[1] The design dimensions and patterns have been discussed in interviews with communication center professionals who have several years of experience in this domain.

[2] The following design patterns have been visualized in form of Petri-nets (Van der Aalst, 1998) with additional communication center specific bitmaps.

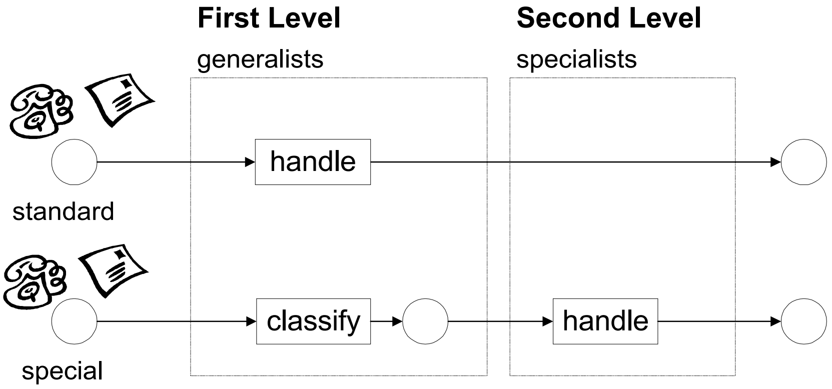

Fig. 2. 2-level pattern

The back-office pattern contrasts with a stronger distinction between synchronous and asynchronous communication. Generalists handle only synchronous requests in the front office whereas specialists classify and handle all asynchronous requests additional to their normal workload of special requests in the back office (Figure 3). This pattern follows the widespread prejudice in practice, that asynchronous communication are more difficult to handle than synchronous communication and should therefore be done by higher qualified employees.

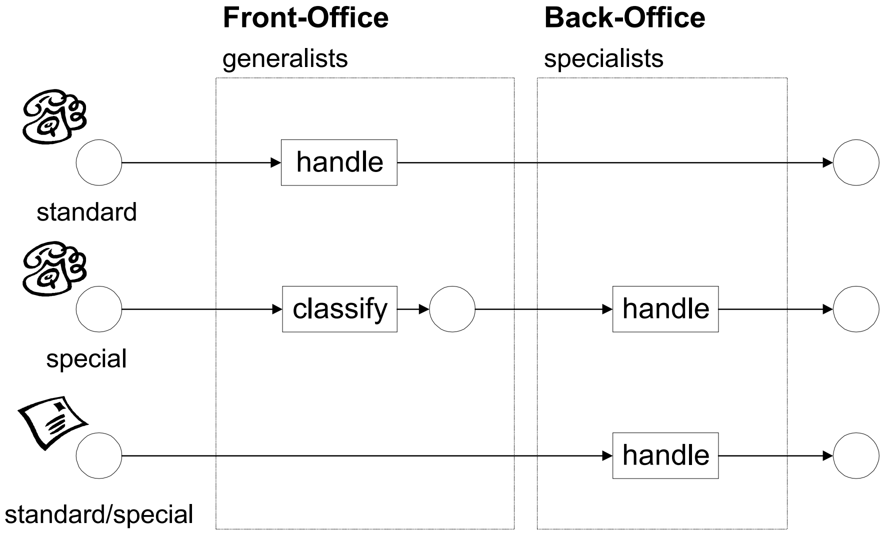

Fig. 3. Back-office pattern

The strongest integration of qualification groups is realized within the 1-level pattern. In this pattern first and second level (or front-office and back-office) will not be distinguished. Generalists and specialists are both able to classify and handle all types of requests (Figure 4). Since most specialists are more expensive than generalists requests are primarily assigned to a free generalist. Only if no generalist is available the request will be assigned to a specialist.

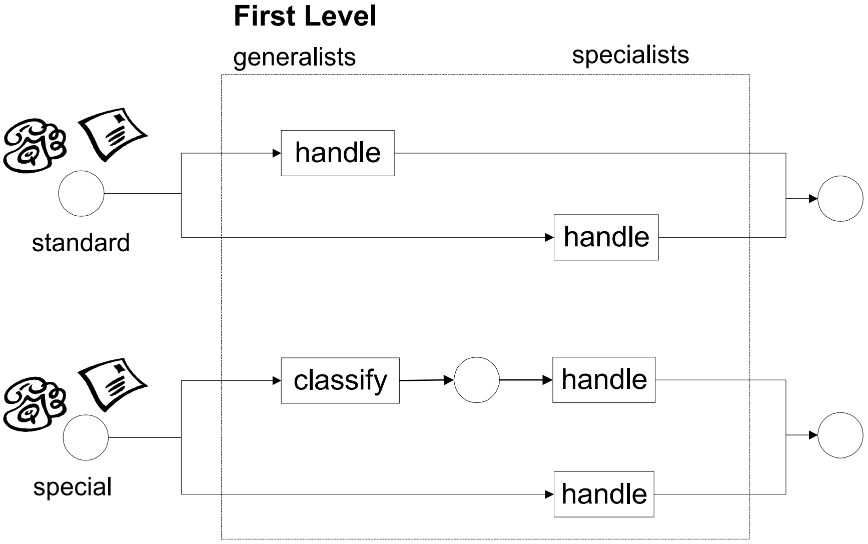

Fig. 4. 1-level pattern

4.1.2 Communication-mixture patterns

The communication-mixture patterns distinguish the (a) integration and (b) separation of communication channels. Integration means that agents handle both synchronous and asynchronous requests. Within the separation of channels different agent groups will be established for synchronous and asynchronous media.

In the preceding paragraph the qualification-mixture patterns 2-level, back-office and 1-level have been presented with integrated communication channels. Synchronous and asynchronous requests will be handled due to their arrival time. If more requests are in the queue synchronous request will be prioritized. The processing of asynchronous requests will not be interrupted by incoming synchronous requests.

In order to achieve separated channels additional agent groups have to be built. In the 2-level pattern two groups “generalists 1” and “specialists 1” will be established for synchronous requests and two additional groups “generalists 2” and “specialists 2” will be built for asynchronous requests. The routing remains the same as shown in Figure 2.

Since in the back-office pattern generalists handle only synchronous requests there is no difference between integrated and separated version for them (Figure 3). An additional group has to be defined in the back office, so that synchronous requests are handled by the group “specialists 1” and asynchronous requests by the group “specialists 2”.

For the separated version of the 1-level pattern four groups (generalists 1, generalists 2, specialists 1, specialists 2) have to be built similar to the 2-level pattern. The basic routing strategy remains the same as shown in Figure 4.

4.2 Basic performance measures and evaluation technique

In the communication center domain many measures are used to evaluate the process performance (Zapf/Heinzl, 2000). Since the goal of a communication center is to handle customer requests it is quite natural that the accessibility is used as an important performance criterion. It is often measured as percentage of lost calls. A high percentage stands for many dissatisfied customers and a poor process performance. In the following the percentage of lost calls is used as basic performance measure for synchronous requests.

Asynchronous requests reach the communication center anyway but are processed with different speed. From the customer point of view a short waiting time is desirable in most of the cases. So the overall waiting time is used as measure for asynchronous requests where a long waiting time indicates a poor performance.

The performance values will be obtained by stochastic discrete event simulation. The single experiments have been performed in the form of multiple terminating simulation runs in order to reflect the nature of a communication center. According to the general rule of Bulgren, 1982, every experiment consists of 30 independent runs.

In order to support the handling of the numerous performed experiments a new database based tool called SimControl has been developed (Zapf, 2001). With SimControl much of the parameterization, initialization, simulation running and data preparation work could be successfully automated. Within SimControl the simulation software ARENA has been used as simulation kernel (Kelton et al., 1998). Some model parts have been created with the call center specific extension Call$im (Systems Modeling Corp., 1996), other parts have been implemented through individual routines.

4.3 Research hypothesis

As basis for the following performance evaluation two research hypothesis will be derived from the literature for qualification-mixture and communication-mixture patterns.

In all of the three qualification-mixture patterns (1-level, 2-level and back-office) customer requests are handled by generalist and specialist groups. But the patterns differ according to (1) the consolidation of the basic activities classify requests and handle requests and (2) the assignment of agents and request types (see Table 3).

(1) According to different coordination approaches the consolidation of sequential dependencies leads to a lower coordination effort and to a better performance (e.g. Thompson, 1967, Kilmann, 1987, Seidmann et al., 1997, Dewan et al., 1997, Buzacott, 1996). This coordination effect is to be expected in our case since different consolidation levels exist among the patterns. (2) Differences in the size of the agent groups lead also to a pooling effect, because bigger groups lead to smaller waiting times and to a better performance (e.g. Biermann et al., 1991, Winston, 1994, Kleinrock, 1976). Since the patterns have different group sizes per request type hypothesis 1 can be derived altogether:

Hypothesis 1: Coordination and pooling effect for the qualification-mixture

The consolidation of the basic activities classify request and handle request and the expansion of the group size leads to a better performance, especially

a) the 1-level pattern is more efficient than the 2-level pattern for all requests;
b) the 1-level pattern is more efficient than the back-office pattern for synchronous standard requests and
c) the back-office pattern is more efficient than the other patterns for all asynchronous requests.

Within the communication-mixture patterns two different approaches exist, the integration and the separation of communication channels. The separation of channels leads to smaller agent groups and therefore the following pooling effect can be expected:

Hypothesis 2: Pooling effect for the communication-mixture

The reduction of agent group members with separated communication channels leads to a poorer performance in comparison to integrated channels.

4.4 Empirical process data

The design patterns have been evaluated for four different lines of business: car rental, financial services, book trade and energy industry (Table 4). The accompanying communication centers perform different tasks but each center has both synchronous and asynchronous communication channels. The information quality is different among the lines of business. For the car rental most information could be obtained with process mining directly from the advanced call distribution system (ACD). Since not all necessary information could be found in the system some values had to be estimated by experts. In the other lines of business estimated and planning data from domain experts have been used since the accompanying projects referred to the set up of new communication centers.

The communication center domain is extremely dynamic. Therefore most of the environmental parameters are non-deterministic with random distribution and will be modeled as random variables. For these parameters average values have been collected which are presented in Table 5.

Incoming requests will be divided up into different request types according to section "request volume" of Table 5. The accompanying arrival rate is modeled as Poisson-distribution (see Law et al., 1991). If calls are in the waiting queue for a long time most of the customers abandon their contact attempt. Some customers recall after a certain amount of time in order to get into contact with an agent of the communication center. This behavior is modeled with the parameters of section "waiting behavior" in Table 5.

line of business	**car rental**	**bank**	**book trade**	**energy industry**
tasks	information, consulting, reservation, complaint, service, emergency help	information, consulting, complaint, service	information, consulting, ordering, complaint, service	information, consulting, complaint, service
communication media	phone, fax, e-mail, letter	phone, fax, e-mail, letter	phone, fax, e-mail, letter	phone, fax, e-mail, letter
information quality	empirical, estimated	estimated	planning	planning

Table 4. Data sources

The calling and handling times will be modeled with the Exponential distribution. This distribution has been chosen because of (1) evidence from the literature (e.g. Kelton et al., 1998, Buzacott, 1996, and Seidmann et al., 1997) and (2) statistical tests with empirical data. For a sample of call times the χ^2-test delivered a P-value of 0,467 and the test of Kolmogorov/Smirnov resulted in a lower bound for the P-value of 0.15. A P-value of more than 0.10 stands for a good correspondence between the distribution and the sample data (Kelton et al., 1998).

	parameter	**car rental**	**bank**	**book trade**	**energy industry**
request volume	request volume per hour	230	264	206	280
	percentage synchronous standard requests	80.26	55.20	88.94	80.96
	percentage synchronous special requests	13.83	24.80	6.06	3.20
	percentage asynchronous standard requests	3.11	13.80	4.68	15.05
	percentage asynchronous special requests	2.80	6.20	0.32	0.79
request handling time	call time synchronous standard requests (min)	3:10	2:06	2:25	3:35
	call time synchronous special requests (min)	2:12	5:46	5:00	5:00
	after call time syn. standard requests (min)	0:12	0:50	1:00	1:55
	after call time syn. special requests (min)	0:21	1:55	1:00	3:00
	handling time asyn. standard requests (min)	3:00	4:00	4:00	7:00
	handling time asyn. special requests (min)	15:00	15:00	10:00	10:00
waiting behavior	waiting tolerance (min)	1:00	1:00	1:00	1:00
	percentage of recalling	75	75	75	75
	time between dial attempts	0:06	0:06	0:06	0:06
classification	manual classification time syn. requests (min)	0:45	0:45	0:45	0:45
	manual classification time asyn. requests (min)	1:00	1:00	1:00	2:00

Table 5. Parameter values

4.5 Experimental design

4.5.1 Model initialization

Before performing the simulation experiments, the different models have to be initialized with the number of agents per agent group. Hereby it has to be kept in mind that even the worst pattern is able to obtain a given efficiency if enough agents are available and in the same way the best pattern obtains a poor performance if not enough agents are available. For a “fair” initialization the following steps have been executed:

1. Obtain performance targets.
2. Determine the minimal amount of agents k_i for every process pattern A_i with which the performance targets can be met (k_i is the total number of agents in the communication center).
3. Determine the agent capacity k as maximum over all process patterns: $k = \max_i(k_i)$.
4. Distribute the overall agent capacity among all agent groups.

As performance targets the following values resulted from interviews with domain specialists: The average lost call rate should be lower than 10%, the average waiting time should be lower than 15 minutes and as a constraint the agent utilization should be less than 80%. The minimal amount of agents has been determined within multiple initialization experiments using the interval bisection method (see Heuser, 1982) for the agent capacity. Figure 5 shows the agent group capacities which resulted from the initializing experiments.

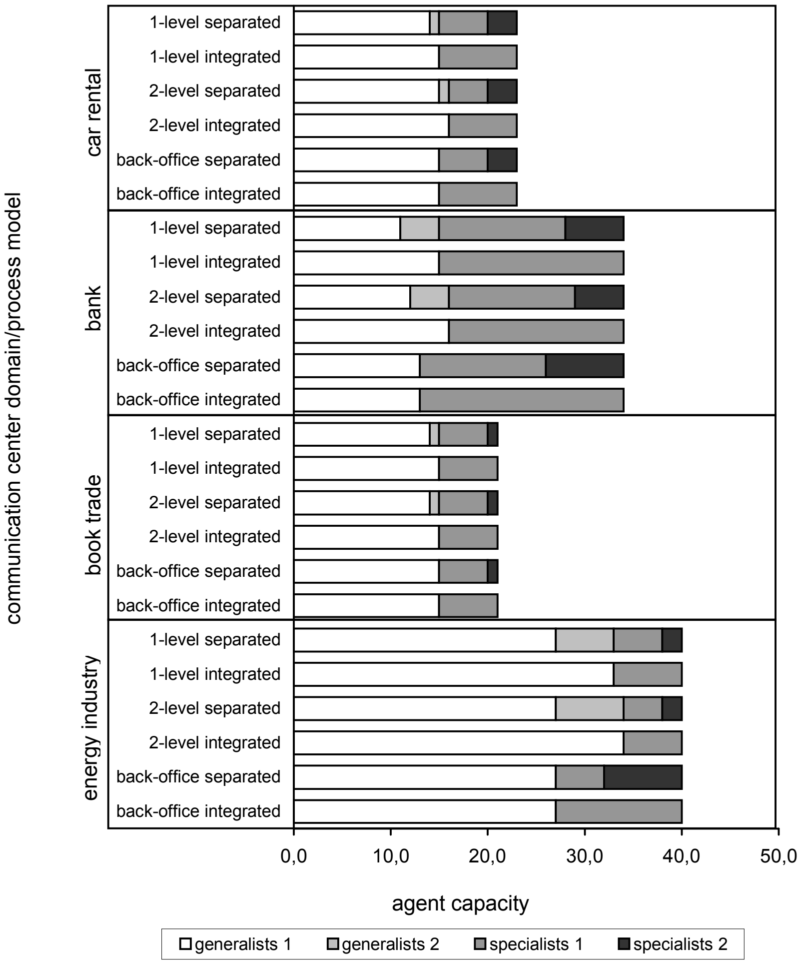

Fig. 5. Agent group capacities

4.5.2 Series of experiments and statistical analysis

First a set of basic experiments with the qualification-mixture and communication mixture pattern has been performed. Within these experiments the efficiency of every pattern from section 4.1 has been determined using the process data from section 4.4.

For qualification-mixture patterns some detailed singular and multiple flexibility analysis has been carried out. Within the singular flexibility analysis we have assumed worst case values for different parameters and changed one parameter after the other within the experimental series. Note that only one parameter has been set to its worst case value per experiment and never multiple parameters have been changed simultaneous. Table 6 shows the parameters and worst case modifications.

Some worst case modifications relate to parameter groups (U_1, U_2, U_4), others to single parameters (U_3, U_5, U_6, U_7). The environments U_1, U_3, U_4, U_5, U_6 and U_7 result from a modification of the average value(s) from Table 5 with different multiplication factors from Table 6. The shape of the statistical distribution is not modified. The modification of the handling time (U_2) results from a special modification of the statistical distribution which is called contamination (see Huber, 1981, and Hampel et al., 1986). Hereby the original distribution is contaminated with *k%* of a special contamination distribution (k = contamination radius). In our case the Exponential distribution is contaminated with a constant distribution which results from doubling the average value of the original distribution.

environment	**description**	**type**	**worst case modification**
U_1 **overload**	increase of request volume	parameter group	multiply with 2
U_2 **long handling**	extension of request handling time	parameter group	100% contamination with the doubled average value
U_3 **long forwarding**	extension of forwarding time	single parameter	multiply with 2
U_4 **agent absence**	reduction of agent availability	parameter group	multiply with 0.8
U_5 **low tolerance**	reduction of waiting tolerance	single parameter	multiply with 0.1
U_6 **increase recalling**	increase of recall percentage	single parameter	multiply with 1.133
U_7 **short recall period**	reduction of time between dial attempts	single parameter	multiply with 0.01

Table 6. Environmental modifications within the singular flexibility analysis

As result of the singular flexibility analysis the parameters with the most impact on the overall performance of the business process have been identified: request volume, handling time, agent availability and average waiting tolerance. These parameters are the basis of a further multiple flexibility analysis in which every possible parameter combination of normal and worst case value has been evaluated (see Table 7).

environment	increase of request volume	contamination of handling time	agent availability	average waiting tolerance (minutes)
U_1	100 %	0 %	100%	1
U_2	100 %	0 %	100%	0.1
U_3	100 %	0 %	80%	1
U_4	100 %	0 %	80%	0.1
U_5	100 %	0.5 %	100%	1
U_6	100 %	0.5 %	100%	0.1
U_7	100 %	0.5 %	80%	1
U_8	100 %	0.5 %	80%	0.1
U_9	150 %	0 %	100%	1
U_{10}	150 %	0 %	100%	0.1
U_{11}	150 %	0 %	80%	1
U_{12}	150 %	0 %	80%	0.1
U_{13}	150 %	0.5 %	100%	1
U_{14}	150 %	0.5 %	100%	0.1
U_{15}	150 %	0.5 %	80%	1
U_{16}	150 %	0.5 %	80%	0.1

Table 7. Environmental modifications within the multiple flexibility analysis

The experimental results have been statistical analyzed with the multiple comparison test from Bonferroni (see Rice, 1988, Kelton et al. 1998). With this test the statistical significance of the observed performance differences between the process patterns has been verified.

4.6 Results and discussion

4.6.1 Coordination and pooling effect for the qualification-mixture

In some patterns the classification and handling activity is consolidated for special requests which leads to performance gains (coordination effect). Apart from that some patterns have a larger group of agents for accepting calls which leads to a more efficient process (pooling effect). Both effects have been stated in hypothesis 1, section 4.3, and evaluated in a series of basic experiments, a singular and a multiple flexibility analysis. The results of this experiments are presented in Table 8.

domain	experimental series[3]	synchronous standard requests			synchronous special requests			asynchronous standard requests			asynchronous special requests		
		1-level	2-level	back-office	1-level	2-level	back-office	1-level	2-level	back-office	1-level	2-level	back-office
car rental	basic[4]	0.00	3.49	5.06	0.00	2.48	3.38	0.00	6.62	0.34	3.34	10.1	0.00
	singular[5]	0.79	0.25	0.10	0.21	0.17	0.11	1.00	0.93	1.00	0.92	0.91	1.00
	multiple	0.39	0.06	0.01	0.03	0.11	0.08	1.00	0.11	1.00	0.61	0.27	1.00
bank	basic	0.00	2.02	1.97	0.00	2.95	2.87	0.00	3.37	0.72	5.58	11.1	0.00
	singular	0.81	0.27	0.35	0.24	0.15	0.05	1.00	0.94	1.00	0.85	0.89	1.00
	multiple	0.39	0.05	0.13	0.09	0.12	0.00	1.00	0.09	0.99	0.56	0.20	0.99
book trade	basic	0.00	6.71	4.72	0.00	3.72	2.68	0.00	12.5	0.00	3.37	12.9	0.00
	singular	0.80	0.06	0.24	0.25	0.08	0.18	1.00	0.88	1.00	0.96	0.87	1.00
	multiple	0.41	0.00	0.09	0.03	0.10	0.12	1.00	0.09	1.00	0.82	0.12	1.00
energy industry	basic	0.00	2.00	4.38	0.35	0.00	2.95	0.00	4.54	2.95	3.56	5.27	0.00
	singular	0.61	0.30	0.13	0.13	0.46	0.24	1.00	0.39	1.00	0.89	0.38	1.00
	multiple	0.21	0.08	0.03	0.01	0.29	0.10	1.00	0.05	0.95	0.57	0.20	0.95

Table 8. Experimental results for qualification-mixture patterns in different domains[6]

Hypothesis 1a: The 1-level pattern is more efficient than the 2-level pattern for all requests.

Hypothesis 1a could be confirmed for synchronous and asynchronous standard requests as well as for synchronous special requests in the domains car rental, bank and book trade within the basic experiments and the singular flexibility analysis. Within the multiple flexibility analysis of all domains the 1-level pattern is inferior to the 2-level pattern. This phenomenon could be explained with a high *specialist occupation* in the 1-level pattern. In overload situations, which are especially examined within the multiple flexibility analysis, standard requests are directly assigned to specialists. This leads to a higher load for specialists and lower capacity for handling special requests. In the energy industry domain hypothesis 1a also does not hold true. Here is the proportion of specialists and with it the possible coordination effect relatively low (see Figure 5). Additionally is the overall handling time for standard requests relatively long (see Table 5) and therefore the specialist occupation effect, which has been described above, leads to a lower specialist capacity and poorer performance in overload situations. For asynchronous requests hypothesis 1a could be verified for all domains except the bank domain within the singular flexibility analysis. This result is primarily caused by poor performance values in the environments overload (U_1) and long handling (U_2) and indicates also the specialist occupation effect.

---

[3] Legend: basic = series of basic experiments, singular = singular flexibility analysis, multiple = multiple flexibility analysis.

[4] Presented are not the absolute but the relative values which have been calculated as difference from the best performance value in the series. Dimension is the average lost call percentage for synchronous requests and the average waiting time in minutes for asynchronous requests.

[5] Presented as aggregate flexibility value without dimension.

[6] The best performance values are shaded gray. If there is no statistical significant difference between the best patterns both values are shaded.

Hypothesis 1b: The 1-level pattern is more efficient than the back-office pattern for synchronous standard requests

Hypothesis 1b has been verified in all experimental series for all domains.

Hypothesis 1c: The back-office pattern is more efficient than the other patterns for all asynchronous requests

The back-office pattern is more efficient than the 2-level pattern for all domains, so hypothesis 1c holds true regarding these both patterns. For asynchronous special requests hypothesis 1c could be verified in all domains also regarding the comparison between 1-level and back-office pattern. In the back-office pattern all asynchronous standard requests are directly accepted and handled by specialists whereas in the 1-level pattern only a part of the asynchronous requests are directly assigned to specialists. Therefore we expected a higher coordination effect and better performance within the back-office pattern. This expectation could not be confirmed within the simulation study, because both patterns achieve generally best performance values for asynchronous standard requests. In the multiple flexibility analysis for the domains bank and energy industry the 1-level pattern is even better than the back-office pattern.

4.6.2 Pooling effect for the communication-mixture

The communication-mixture patterns comprise separated and integrated communication channels. In the case of separated channels the number of agents per group is generally smaller compared with integrated channels. Therefore the separation of channels leads to a smaller pooling effect and is expected to be less efficient than the channel integration. This effect has been stated in hypothesis 2 and evaluated within a series of basic experiments.

Hypothesis 2: Separated communication channels lead to a poorer performance than integrated channels

Hypothesis 2 could be verified for the 1-level and 2-level patterns in all domains (see Figure 6). In the back-office pattern there are no significant differences for synchronous standard requests. These requests are handled completely by generalist agents who are exclusively responsible for synchronous requests and therefore not affected by the separation of communication channels.

For synchronous special requests there appear efficiency gains for separated channels in the car and book domain, which has not been expected before. This could be explained with a competition effect which operates contrarily to the pooling effect. In the case of integrated communication channels an agent handles asynchronous requests as soon as no synchronous request exists in the queue. During the processing time the agent is occupied and cannot accept a new synchronous request. Therefore synchronous and asynchronous requests compete for the same agents and disadvantages arise for accepting synchronous requests.

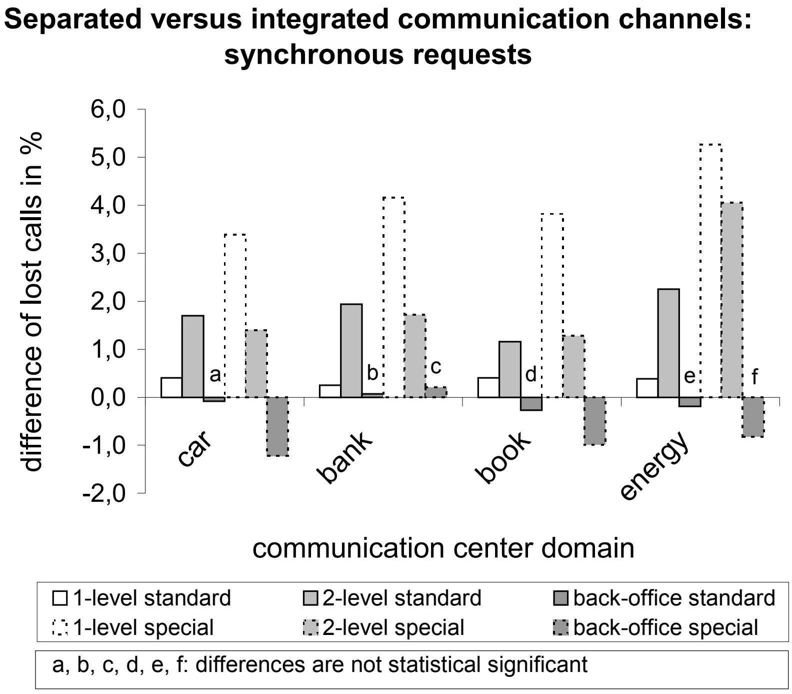

Fig. 6. Differences between separated and integrated communication channels for synchronous requests with different qualification-mixtures, request types and domains

For asynchronous requests the separation of communication channels leads to a slower processing speed (see Figure 7).

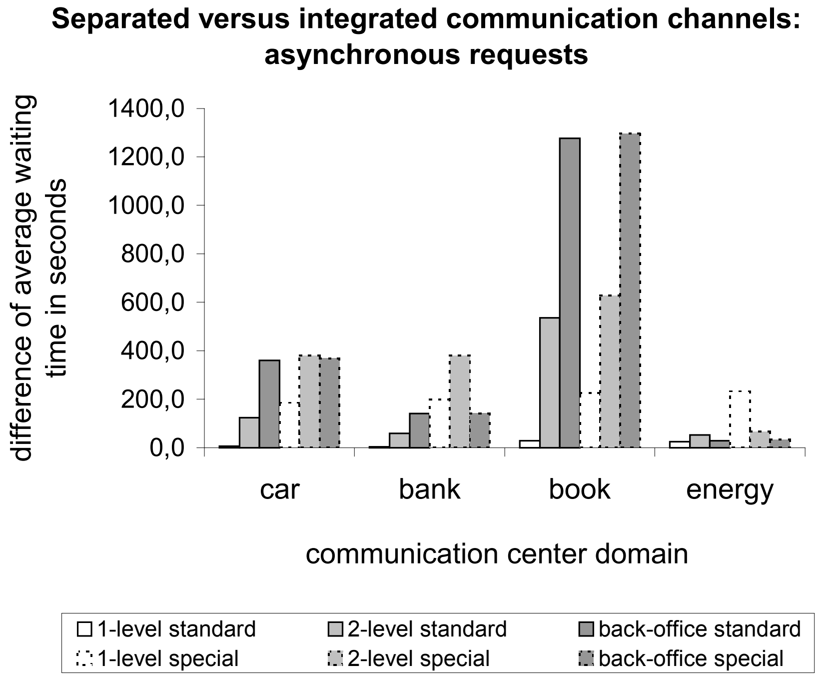

Fig. 7. Differences between separated and integrated communication channels for asynchronous requests with different qualification-mixtures, request types and domains

This confirms hypothesis 2. In the energy domain the pooling effect is unincisive and very strong in the book trade domain. This can be explained as follows: In the book domain there are only 1 or 2 agents handling asynchronous requests. With separated channels only 1 agent is available per "group" and therefore an extremely bad performance can be expected in overload situations. In the energy domain 8 to 9 agents are processing asynchronous requests and the separation results in bigger groups. Here the pooling effect remains also for separated channels and the loss of performance is not so strong as in the book trade domain.

5 Conclusion and further research

In this paper we presented the pattern-driven process design approach. It comprises a procedure for evaluating and selecting a suitable process pattern for a given purpose and a technique for measuring the flexibility of business processes. This approach has been introduced and applied within an extensive evaluation of communication center patterns on the basis of empirical data from this domain.

Three patterns have been evaluated regarding the qualification-mixture in a communication center: the 1-level, 2-level and back-office pattern. It could be verified that the consolidation of tasks in the 1-level pattern leads to efficiency gains (coordination effect). But for some request types a reverse effect could be identified: performance losses are caused for special requests by a higher specialist occupation with standard requests. In most of the results the 2-level pattern was the least efficient alternative whereas this pattern is frequently used in practice. The back-office pattern has weaknesses in accepting synchronous requests but advantages in handling asynchronous requests. The 1-level pattern is for standard requests the most efficient pattern, for special requests some drawbacks have to be kept in mind.

Regarding the communication-mixture two alternative designs have been evaluated: the integration and the separation of communication channels. The integration leads to larger groups and therefore to a higher efficiency (pooling effect). Surprisingly this effect could not be verified for synchronous special requests within the back-office pattern. This has been explained with a competition effect because of the occupation of agents with asynchronous requests.

With the application of flexibility evaluation the following insights have been received: (1) Important differences between the patterns have been made clear only within the explicit analysis of process flexibility. The evaluation without concerning flexibility issues can easily lead to wrong design decisions. (2)There is no flexible or non-flexible process. The process flexibility has to be related to specific process objects like request types. Frequently a process reacts different flexible for different process objects. (3) For the process flexibility not only the availability of slack resources but also the potential to activate these resources is critical.

Concerning our approach some limits have to be kept in mind: (1) We focus on quantitative performance measures and disregard qualitative measures. Qualitative measures like conversation quality or customer satisfaction are very import for the overall performance but if no agent is accessible no conversation takes place and the customer could not be satisfied at all. So quantitative measures are the basis but not sufficient for an overall evaluation of organizational designs. Qualitative measures have to be included in further extensions of the approach. (2) The utilization of simulation does only allow statements regarding the examined parameter constellations and no general deductions as with queueing theory. But without simulation the real life patterns could not have been evaluated. We also performed numerous experiments under different parameter constellations in order to get a wide base for our deductions. (3) Within the flexibility analysis we used subjective threshold values in order to reduce the analysis complexity. The impact of threshold values on the aggregate flexible values is another research topic for the future.

The following research directions have been identified: (1) Based on the presented results further communication center patterns should be identified and analyzed, like sophisticated skill based routing concepts, overflow strategies, call routing between different locations, integration of outbound activities and integration of outsourcing provider. (2) Furthermore relationships between successive requests (customer history) and relationships between requests of different communication channels should be included into the simulation models in order to evaluate socio-technical theories like the Media Richness Theory (Daft et al., 1984) or Social Presence Theory (Short et al., 1976). (3) The approach of measuring process flexibility should be testes in other domains. The ability to use other base performance measures has to be verified and the influence of threshold values has to be evaluated in detail. (4) A strong research direction is the automatic process design and improvement. With the workflow mining techniques it will be possible to identify process patterns faster and more reliable than traditional process modeling by a business consultant. These patterns can be evaluated and it may be possible to automatically generate an improved version of the business process. This would be an enormous progress in the design of organizational processes.

6 References

Becker, J.; Schütte, R., 1996, Handelsinformationssysteme (Landsberg).

Becker, J.; Kugeler, M. and Rosemann, M., 2001, Prozessmanagement. Ein Leitfaden zur prozessorientierten Organisationsgestaltung (Berlin et al.).

Bierman, H.; Bonini, C.P. and Hausman, W.H., 1991, Quantitative Analysis for Business Decisions (Boston).

Bulgren, W.G., 1982, Discrete System Simulation (Englewood Cliffs).

Buzacott, J.A., 1996, Commonalities in Reengineered Business Processes: Models and Issues, Management Science 42 (5), 768-782.

Crowston, K., 1997, A Coordination Theory Approach to Organizational Process Design, Organization Science 8 (2), 157-175.

Daft, R.L. and Lengel, R.H., 1984, Information Richness: A new Approach to Managerial Behavior and Organization Design, Research in Organization Behavior, 6, 191-233.

Dewan, R.; Seidmann, A. and Walter, Z., 1997, Workflow Redesign through Consolidation in Information-Intensive Business Processes, Proceedings of the ICIS 1997, 285-295.

Hammer, M. and Champy, J., 1993, Reengineering the Corporation (New York).

Hampel, F.R.; Ronchetti, E.M. and Rousseuw, P.J., 1986, Robust Statistics (New York et al.).

Hanssmann, F., 1987, Einführung in die Systemforschung (München et al.).

Heuser, H., 1982, Lehrbuch der Analysis, Band 1 (Stuttgart).

Huber, P.J., 1981, Robust Statistics (New York et al.).

Jacob, H., 1974, Unsicherheit und Flexibilität – Zur Theorie der Planung bei Unsicherheit, Zeitschrift für Betriebswirtschaft 44 (5-8), 299-326, 403-448 and 505-525.

Kelton, W.D.; Sadowski, R.P. and Sadowski, D.A., 1998, Simulation with ARENA (Boston et al.).

Kilman, R.H., 1987, Beyond the Quick Fix: managing Five Tracks to Organizational Success (London et al.).

Kleinrock, L., 1976, Queueing Systems (New York et al.).

Kühn, M. and Schneeweiss, C., 1990, Zur Definition und gegenseitigen Abgrenzung der Begriffe Flexibilität, Elastizität und Robustheit, Schmalenbachs Zeitschrift fuer betriebswirtschaftliche Forschung (5), 378-395.

Malone, T.W.; Crowston, K.; Lee, J.; Pentland, B.; Dellarocas, C.; Wyner, G.; Quimby, J.; Osborn, C.; Bernstein, A.; Herman, G. and Klein, M., 1999, Tool for Inventing Organisations: Toward a Handbook of Organizational Process, Management Science, 45 (3), 425-443.

Rice, J.A., 1988, Mathematical Statistics and Data Analysis (Pacific Grove).

Scheer, A., 1994, Business Process Engineering - Reference Models for Industrial Companies (Berlin et al.).

Schütte, R., 1998, Grundsätze ordnungsmäßiger Referenzmodellierung (Wiesbaden).

Seidmann, A. and Sundararajan, A., 1997, The Effects of Task and Information Asymmetry on Business Process Redesign, International Journal of Production Economics 50 (2-3), 117-128.

Short, J.; Williams, E. and Christie, B., 1976, The Social Psychology of Telecommunications (London).

Systems Modeling Corp., 1996, Call$im Template User's Guide (Sewickley).

Thompson, J.D., 1967, Organizations in Action: Social Science Bases of Administrative Theory (New York et al.).

Van der Aalst, W.M.P., 1998, The Application of Petri Nets to Workflow Management, Journal of Circuits, Systems and Computers 8 (1), 21-66.

Van der Aalst, W.M.P., Desel, J. and Oberweis, A., 2000, Business Process Management: Models, Techniques, and Empirical Studies, LNCS 1806 (Berlin et al.).

Van der Aalst, W.M.P. and van Dongen, B.F., 2002, Discovering Workflow Performance Models from Timed Logs. in: Han, Y. ; Tai, S. and Wikarski, D., eds., International Conference on Engineering and Deployment of Cooperative Information Systems (EDCIS 2002), LNCS 2480, 45-63.

Van der Aalst, W.M.P.; ter Hofstede, A.H.M.; Kiepuszewski, B. and Barros, A.P., 2003, Workflow Patterns, Distributed and Parallel Databases (to appear).

Winston, W.L., 1994, Operations Research Applications and Algorithms (Delmont).

Zapf, M. and Heinzl, A., 2000, Evaluation of Generic Process Design Patterns: An Experimental Study, in van der Aalst, W.M.P.; Desel, J. and Oberweis, A., eds., Business Process Management: Models, Techniques, and Empirical Studies, LNCS 1806 (Berlin et al.).

Zapf, M., 2001, Gestaltung flexibler Kundeninteraktionsprozesse im Communication Center: Theoretische Grundlagen und experimentelle Analyse, Dissertation (University of Bayreuth).

Zapf, M. and Storch, K., 2001, Making Simulation Work for the Organizational Design of Communication Centers: Challenges and Practical Experience, in: Proceedings of the 2001 Summer Computer Simulation Conference (Orlando).