Feedback about the study: "What does a successful evaluation project need?" © 28.09.2004

INTRODUCTION

With this text I would like to provide you with the promised feedback about my study "What does a successful evaluation project need?".
You can learn more (status quo, information, results, and questionnaires) about this study at: http://www-user.rhrk.uni-kl.de/~balzer/eval-success.html.

Again thank you very much to all who have participated at my study!

Now it is time to have a look at first results:

My study looks at the very basics when doing high-quality evaluation projects and asks one question which appears to be very simple:
"What does a successful evaluation project need?"
There is much literature around which focuses on this question. You can find "how-to books", a nearly uncountable number of evaluation theories and models, and the discussion about evaluation standards deals with this problem, too.
When reading such texts one can learn a lot about how to manage an evaluation project.
Among others one topic seems to be very important in most texts: The importance to consider the needs of groups which are differently involved in an evaluation project. But it is interesting to see that these different experiences, views and thoughts of different groups are nearly not present in empirically focused research.
But knowing if needs of different groups differ from each other might be very important for practical evaluation work.
Therefore, three main questions are addressed here:

- What conditions for successful evaluation-projects can be identified?
- Do groups which are involved differently in an evaluation project judge conditions in a different way?
- What about the practical evaluation work regarding these conditions?

© by Lars Balzer, Centre for Educational Research, University of Koblenz Landau, Campus Landau, Bürgerstr. 23, 76829 Landau, Germany, www.zepf.uni-landau.de, balzer@rhrk.uni-kl.de

METHODS

To answer these questions an internet-survey was implemented where 103 conditions of successful evaluation projects have been presented. These conditions have mainly been taken from a recent expert-opinion survey (Balzer, 2004), where more than 400 evaluation experts from different fields and domains had worked out about 100 conditions.
In this study evaluation experts have been asked to judge these conditions regarding the two following questions:

Please give your opinion in consideration to the two following questions:
- How do you assess these conditions?
(1 = without this no successful evaluation project is possible
2 = very important condition
3 = important condition
4 = relevant condition, but you can manage without it
5 = unimportant condition
6 = this condition is even counterproductive)
- According to your experience: What is the current evaluation practice regarding these conditions?
(1 = receives too much consideration
2 = receives appropriate consideration
3 = receives too little consideration)

Evaluation experts in this study are defined in a very broad sense:
Evaluation experts are people,

- who are stakeholders within an evaluation project and act as experts for their own experiences, wishes, needs and thoughts
- who are outstanding persons in the academic and/or practical field of evaluation (not necessarily involved in an evaluation project, but with important experience in the field)
- who do evaluation projects themselves

Experts have been asked to participate in this study in different ways:

- A personal mail or letter was sent directly to experts
- Information about the study (and a link to the project WWW-site; still online at: http://www-user.rhrk.uni-kl.de/~balzer/eval-success.html) has been disseminated in evaluation-newsgroups and evaluation mailinglists.
- National and international evaluation associations have been contacted and informed.
- Participants of the study have been asked to spread the information and/or to give names of others who might be interested in the study. Those have been contacted, too.
- Several (internet)-searches have given information about further institutions and persons which have been contacted then, too.

SAMPLE

Altogether 442 experts participated at the study. Mainly they have used the internet-questionnaire, sporadically the paper-pencil version was used.
Following up some characteristics of the sample are presented:
245 (55,9%) are female, 193 (44,1%) are male and 4 have not given information about their gender.
Age has a range from 22 to 82 years, with a mean of 43,9 years (SD = 11,3).
People from nearly 50 different countries have participated, with an emphasis on the USA with a bit more than 50% (see table 1).

Table 1:
Country of origin

	N	percent
USA	218	52,0
Australia	21	5,0
Canada	20	4,8
England	16	3,8
Sweden	15	3,6
Scotland	14	3,3
Finland	12	2,9
Ireland	12	2,9
Germany	8	1,9
Brazil	7	1,7
Israel	7	1,7
France	6	1,4
Switzerland	6	1,4
South Africa	5	1,2
Denmark	4	1,0
Wales	4	1,0
Belgium	3	,7
Austria	2	,5
Colombia	2	,5
India	2	,5
Mexico	2	,5
Netherlands	2	,5
Norway	2	,5
Pakistan	2	,5
Russia	2	,5
Tanzania	2	,5
Cameroon	2	,5
Puerto Rico	2	,5
Kenya	2	,5
Cyprus	1	,2
Greece	1	,2
Malaysia	1	,2

table 1 (continued)	N	percent
New Zealand	1	,2
Portugal	1	,2
Spain	1	,2
Turkey	1	,2
Ukraine	1	,2
Lebanon	1	,2
Madagascar	1	,2
Moldavia	1	,2
Uganda	1	,2
Senegal	1	,2
Tunisia	1	,2
Ghana	1	,2
Ethiopia	1	,2
Nigeria	1	,2
total	419	100,0
no answer	23	

Looking at the highest degree or academic title the experts have achieved one can see that most of the participants have an academic background (see table 2). The huge number of other academic degrees represents mainly Master degrees in different domains.

Table 2:
Highest degree achieved

	N	percent
none	0	0,0
I go to school	1	0,2
graduation	5	1,1
I serve my apprenticeship	0	0,0
I have finished my apprenticeship	0	0,0
foreman	0	0,0
other non-academic degree	0	0,0
I go to university	3	0,7
degree at polytechnic/ vocational college or similar	2	0,5
university degree	143	32,8
PhD	170	39,0
private lecturer / PD	5	1,1
professor	20	4,6
other academic degree	87	20,0
total	436	100
no answer	6	

Most people with academic background come from the social sciences (see table 3). But about 30% of the participants have difficulties to classify themselves within the provided categories. Having a closer look at the category “other field” one can see many different fields with an emphasis on education and health.

Table 3:
Working field of the academics

	N	percent
social sciences	222	53,9
linguistics and cultural sciences	4	1,0
medicine	16	3,9
mathematics and natural sciences	10	2,4
engineering sciences	2	0,5
economic sciences	14	3,4
sciences of sport	0	0,0
law	1	0,2
political sciences	19	4,6
other field	124	30,1
total	412	100,0
no answer	15	

9 persons do not have an academic background and most of them come from the field of education.

The institutions and organizations, where the evaluation experts are working at, are mainly non-profit (293 = 74,4% non-profit, 101 = 25,6% profit, 48 no answer) and public (234 = 59,5% public, 159 = 40,5% private, 49 no answer). Nearly 50% are universities or other research orientated institutions (185 = 45,0% university/research, 98 = 23,8% authorities/ government, 69 = 16,8% private enterprise/industry, 59 = 14,4% other, 31 no answer).

There are various domains in which the experts are in touch with evaluation.
Nearly ¾ deal with evaluation methods and nearly half of them deals with evaluation in school education. Furthermore, evaluation in academic education, evaluation theory and evaluation in the field of medicine and health care seems to be important (see table 4).

Table 4:
Domain in which participants are in touch with evaluation (more than one can be chosen)

	N	percent
medicine and health care	124	28,1
school education	201	45,5
special needs education	58	13,1
vocational education	62	14,0
academic education	168	38,0
further education	57	12,9
politics	56	12,7
economy	42	9,5
trade and engineering	11	2,5
computing, telecommunication, internet	58	13,1
law	13	2,9
military	7	1,6

table 4 (continued)	N	percent
evaluation methods	320	72,4
theory of evaluation	160	36,2
other domain	149	33,7

Looking into details evaluation methods and evaluation in school education seem to be most important. In each case more than 20 percent say that one of this is his/her most important domain in which he/she is in touch with evaluation. About every seventh opts for evaluation in medicine and health care, and about every tenth for evaluation in academic education and evaluation in another domain. Remaining domains are only sporadically nominated to be most important (see table 5).

Table 5:
Most important domain in which participants are in touch with evaluation (only one can be chosen)

	N	percent
medicine and health care	49	15,6
school education	64	20,3
special needs education	6	1,9
vocational education	7	2,2
academic education	34	10,8
further education	10	3,2
politics	10	3,2
economy	10	3,2
trade and engineering	3	1,0
computing, telecommunication, internet	6	1,9
law	4	1,3
military	0	0,0
evaluation methods	69	21,9
theory of evaluation	9	2,9
other domain	34	10,8
total	315	100,0
no answer	127	

Because this study is an expert-opinion survey it is not surprising that much evaluation expertise can be found. Table 6 shows, that practical orientated evaluation work is even more important for this sample than theoretically orientated evaluation work. The difference is statistically significant ($t = -14{,}36$; $df = 424$; $p = 0{,}000$) and has a high effect size ($\omega^2 = 0{,}326$).

Table 6:
Evaluation expertise

	work with evaluation			
	theoretically orientated		practically orientated	
	N	percent	N	percent
highest expertise	38	8,6	135	31,0
high expertise	155	35,1	189	43,4
medium expertise	179	40,5	96	22,1
low expertise	51	11,5	15	3,4
no expertise	4	0,9	0	0,0
total	427	100,0	435	100,0
no answer	15		7	

Looking at the links to the field of evaluation table 7 shows that most experts use results of evaluation projects, do evaluation projects themselves, act as experts or are consultants within evaluation projects. More rarely experts educate in the field of evaluation, are interviewees in evaluation projects or are clients of evaluation projects.

Table 7:
Personal link to the field of evaluation

	I do evaluation projects myself		I am a client of evaluation projects	
	N	percent	N	percent
last year	289	67,5	77	23,0
in the last five years	82	19,2	69	20,6
longer ago than 5 years	40	9,3	43	12,8
never	17	4,0	146	43,6
total	428	100,0	335	100,0
no answer	14		107	
	I am an interviewee within an evaluation project		I am using evaluation results	
	N	percent	N	percent
last year	91	27,5	269	71,7
in the last five years	91	27,5	61	16,3
longer ago than 5 years	39	11,8	30	8,0
never	110	33,2	15	4,0
total	331	100,0	375	100,0
no answer	111		67	

table 7 (continued)				
	I am a consultant within evaluation projects		I educate in the field of evaluation	
	N	percent	N	percent
last year	232	60,7	176	49,2
in the last five years	69	18,1	65	18,2
longer ago than 5 years	30	7,9	27	7,5
never	51	13,4	90	25,1
total	382	100,0	358	100,0
no answer	60		84	
	I act as an expert within evaluation projects		other link to evaluation	
	N	percent	N	percent
last year	233	61,6	44	39,3
in the last five years	60	15,9	6	5,4
longer ago than 5 years	29	7,7	6	5,4
never	56	14,8	56	50,0
total	378	100,0	112	100,0
no answer	64		330	

Asked for the evaluation object which the experts are mostly experienced with more than half of the experts mentioned programmes (see table 8).

Table 8:
Experiences with different evaluation objects

	N	percent
products	8	1,8
people	23	5,2
programmes	245	55,7
interventions/treatments	75	17,0
institutions/systems	45	10,2
politics	12	2,7
I do not have to emphasize one of them	32	7,3
total	440	100,0
no answer	2	

RESULTS

On the one hand the main interest of this study is to judge the importance of conditions of successful evaluation projects in the view of all experts as well as in the view of different expert subgroups.
On the other hand this study should show how the current evaluation practise is for all these conditions.

Following up one will find results which are addressed to these ideas. To identify different groups of evaluation experts information presented in table 7 was taken to build subgroups. Out of 442 experts there are 189 persons who are (among others) clients of evaluation projects.
232 are evaluators who are no clients. 4 are interviewees in evaluation projects without being clients or evaluators. 17 belong to other groups.
Taking the number of persons within one group into account following up the clients and the evaluators are analyzed in detail. Unfortunately other groups are too small to be analyzed.

RESULTS – IMPORTANCE OF CONDITIONS

First of all the importance of the conditions for successful evaluation projects assessed by the experts is analyzed in detail. As mentioned above the experts had 6 possibilities to rank each condition (1 = without this no successful evaluation project is possible; 2 = very important condition; 3 = important condition; 4 = relevant condition, but you can manage without it; 5 = unimportant condition; 6 = this condition is even counterproductive).
The following tables show data of all 103 conditions sorted by 7 categories ("evaluation basics", "characteristics of the client", "characteristics of the evaluator", "characteristics of participants within evaluation projects", "characteristics of the outer field", "realisation of the evaluation project" and "realisation of the evaluation project").
For this purpose sequential ranks have been built for each condition (lower rank = more important). The mean ranks (MR) are shown for all 442 experts (all), for 189 clients (client) and 232 evaluators (eval.). Numbers can differ from one condition to the other because of missing data. Differences between evaluators and clients are statistically tested (p and ω^2 if necessary). First of all the seven tables are presented. After that the results are commented.

Table 9:
Category "Evaluation basics"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
1- clear definition of the object of the evaluation	1,38	0,67	1,38	0,68	1,39	0,67	,817	
2- clear and realistic evaluation objectives and questions	1,51	0,70	1,49	0,74	1,54	0,66	,190	
3- clear definition of the evaluation context	1,81	0,80	1,82	0,80	1,83	0,82	,982	
4- clear definition of evaluation participants	2,06	0,95	2,07	0,94	2,09	0,95	,866	
5- clear basis of evaluation and assessment indicators	2,08	0,95	2,10	1,00	2,10	0,90	,674	
6- formal agreement of evaluation objectives and questions	2,10	0,99	2,16	1,00	2,08	0,98	,407	
7- evaluation objectives take interests of different stakeholders into consideration	2,14	0,97	2,13	1,02	2,17	0,92	,358	
8- similar evaluation understanding between client and evaluator	2,14	1,07	2,17	1,07	2,14	1,07	,775	
9- precise description of information sources	2,32	0,92	2,35	0,91	2,33	0,93	,799	
10- usage of evaluation results is clarified in advance	2,38	1,03	2,42	0,97	2,40	1,08	,600	

table 9 (continued)	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
11- tendering procedure of the evaluation project	2,53	1,07	2,39	0,98	2,66	1,12	,047	,009
12- no hidden evaluation objectives	2,57	1,29	2,50	1,20	2,70	1,36	,279	
13- evaluation is not used in the sense of controlling	2,72	1,31	2,66	1,23	2,83	1,33	,210	
14- (theoretical) foundation of the evaluation object	2,75	1,00	2,84	0,91	2,72	1,02	,221	
15- establishment of an evaluation culture	2,75	1,13	2,72	1,11	2,84	1,14	,311	
16- service orientation of the evaluation	2,76	1,11	2,84	1,08	2,78	1,14	,457	
17- enough time between contract and the beginning of the evaluation project	2,77	1,10	2,84	1,08	2,76	1,12	,464	
18- reasonable cost-value ratio	2,86	0,97	2,89	0,95	2,89	0,97	,956	
19- evaluation is embedded in total quality management process	3,15	1,22	3,21	1,22	3,12	1,21	,545	
20- evaluation as an independent project, no appendix	3,38	1,32	3,46	1,34	3,37	1,26	,422	

Table 10:
Category "Characteristics of the client"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
21- cooperation and committed participation of the client	1,88	0,86	1,89	0,85	1,91	0,88	,827	
22- openness of the client towards the evaluation and its results	2,04	0,92	1,97	0,91	2,14	0,92	,055	
23- client is willing to change things	2,40	1,07	2,41	1,09	2,46	1,07	,572	
24- client has the power to realize potential changes	2,66	1,09	2,57	1,10	2,76	1,09	,110	
25- client has expertise in the field of evaluation	4,11	0,97	4,11	1,01	4,15	0,91	,991	
26- client takes a checking and controlling function within the evaluation project	4,35	1,47	4,36	1,44	4,42	1,47	,678	
27- client keeps out of the evaluation project	4,52	1,50	4,62	1,50	4,42	1,50	,178	
28- client is independent from the decision makers	4,56	1,46	4,55	1,55	4,57	1,34	,633	

Table 11:
Category "Characteristics of the evaluator"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
29- incorruptibility of the evaluator	1,62	0,82	1,48	0,71	1,76	0,91	,006	,020
30- acceptance and credibility of the evaluator	1,78	0,76	1,72	0,72	1,85	0,79	,167	
31- high evaluation-expertise of the evaluator	1,98	0,77	1,97	0,75	2,02	0,77	,560	
32- cooperation between evaluator and internal staff	2,02	0,90	2,07	0,94	1,98	0,85	,449	
33- objectivity and neutrality of the evaluator	2,07	1,11	1,91	1,07	2,24	1,13	,003	,025
34- motivated evaluator	2,09	0,84	2,08	0,78	2,12	0,91	,898	
35- self-reflexion capability of the evaluator	2,19	0,89	2,22	0,95	2,19	0,85	,900	
36- independence of the evaluator	2,26	1,15	2,24	1,15	2,31	1,15	,564	
37- high social competence of the evaluator	2,27	0,91	2,20	0,93	2,34	0,89	,196	
38- evaluator has expertise in the field where the evaluation takes place	2,73	1,09	2,63	1,07	2,79	1,06	,148	
39- exchange of experiences with other evaluators	2,84	0,96	2,78	0,90	2,92	1,02	,310	
40- heterogeneous evaluation team	3,23	1,26	3,10	1,27	3,34	1,25	,076	

Table 12:
Category "Characteristics of participants within evaluation projects"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
41- cooperation and committed participation of all participants	2,34	0,97	2,34	0,96	2,37	0,97	,710	
42- participants accept the evaluation plan	2,40	1,00	2,25	0,88	2,56	1,08	,014	,016
43- openness of the participants towards evaluation results	2,45	0,95	2,38	0,89	2,58	0,96	,052	
44- voluntary participation of the participants	2,62	1,17	2,61	1,08	2,68	1,21	,869	
45- the participants are accepted by the persons responsible for the project	2,78	1,18	2,87	1,13	2,77	1,20	,430	
46- volition of participants to change something	2,84	1,10	2,70	0,98	3,03	1,18	,014	,017
47- professional competence, familiarity of participants with the evaluation object	2,95	1,09	2,82	0,98	3,06	1,14	,069	
48- the participants have experience with internal evaluation	4,08	0,99	4,08	0,93	4,13	1,03	,311	
49- the participants have evaluation expertise	4,41	0,94	4,46	0,89	4,41	0,95	,786	
50 -participants take part because of client's enforcement-power	5,00	1,29	5,01	1,33	5,02	1,24	,761	

Table 13:
Category "Characteristics of the outer field"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
51- favourable general political conditions	3,14	1,08	3,19	1,08	3,10	1,11	,692	
52- interest of the wider environment in the evaluation	3,33	1,03	3,35	0,99	3,37	1,04	,768	

Table 14:
Category "Realisation of the evaluation project"

	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
53- clear and adequate evaluation design	1,66	0,88	1,70	0,99	1,65	0,77	,810	
54- consideration of ethics	1,79	0,82	1,73	0,78	1,85	0,86	,245	
55- correct methodological procedures of evaluation	1,89	0,83	1,92	0,81	1,91	0,85	,814	
56- clear responsibility assignment of the persons responsible for the project	2,05	0,90	2,00	0,85	2,13	0,94	,243	
57- clear identification of all stakeholders	2,08	0,88	2,07	0,80	2,13	0,96	,809	
58- consideration of legal basis	2,08	0,93	1,97	0,89	2,18	0,96	,061	
59- consideration of relevant evaluation standards	2,11	0,85	2,01	0,85	2,24	0,86	,027	,013
60- existence of sufficient resources	2,14	0,88	2,10	0,84	2,15	0,90	,716	
61- attention to data protection	2,14	0,91	2,12	0,96	2,18	0,89	,391	
62- adequate involvement of all participants	2,15	0,88	2,08	0,83	2,23	0,92	,178	
63- transparency of evaluation process	2,22	0,96	2,16	1,00	2,30	0,95	,190	
64- data is accessable with justifiable effort	2,23	0,86	2,16	0,82	2,31	0,91	,252	
65- flexibility during the data collection	2,26	0,86	2,26	0,85	2,29	0,88	,796	
66- quantitative as well as qualitative procedure	2,34	1,09	2,26	1,07	2,42	1,09	,146	

table 14 (continued)	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
67- ongoing discussion about the evaluation process	2,48	0,98	2,35	0,96	2,65	0,98	,007	,021
68- process-orientated procedure	2,60	1,09	2,60	1,05	2,63	1,12	,684	
69- evaluation is practice-orientated, so that it disturbs everyday life as little as possible	2,61	1,08	2,64	1,06	2,65	1,11	,956	
70- consensus regarding evaluation procedure between as many stakeholders as possible	2,76	1,05	2,69	1,11	2,89	0,97	,036	,012
71- consideration to hidden goals	2,82	1,30	2,87	1,33	2,76	1,22	,679	
72- cooperation between all stakeholders	2,86	1,08	2,85	1,06	2,89	1,08	,569	
73- responsibility of realization belongs to evaluator	2,91	1,41	2,97	1,46	2,82	1,32	,436	
74- permanent control of the evaluation process	3,24	1,24	3,34	1,22	3,19	1,20	,413	
75- evaluation project has an advisary board	3,45	1,09	3,47	1,10	3,41	1,07	,610	

Table 15:
Category "Results of the evaluation project"

	all(442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
76- results are complete and fair	1,71	0,67	1,68	0,67	1,74	0,69	,486	
77- report transparently describes the object of the evaluation, its context, objectives, procedure and results	1,79	0,81	1,76	0,81	1,85	0,83	,319	
78- correct methodological procedure of data analysis	1,81	0,85	1,80	0,85	1,82	0,83	,685	
79- comprehensible, receiver adequate report of results	1,89	0,83	1,80	0,87	1,97	0,80	,037	,012
80- sensitive, confidential, factual handling of results	1,89	0,93	1,77	0,82	2,00	0,99	,066	
81- (potentially) helpful and constructive results	1,97	0,89	2,09	0,98	1,90	0,80	,164	
82- reasons for interpretations are explicitly given	2,00	0,99	1,86	0,87	2,16	1,10	,032	,013
83- discussions about results	2,03	0,89	1,97	0,90	2,11	0,90	,167	
84- report timeliness	2,16	0,90	2,14	0,93	2,17	0,88	,680	
85- report contains no rash conclusions	2,16	1,28	2,14	1,36	2,22	1,23	,238	
86- transparency of using the results	2,33	0,95	2,29	0,98	2,42	0,92	,137	
87- report strictly segregates results, interpretations and recommendations	2,39	1,40	2,28	1,44	2,48	1,35	,086	
88- disclosure of results to all stakeholders	2,43	1,05	2,39	1,07	2,54	1,04	,173	
89- report contains interpretations by the evaluator	2,51	1,03	2,44	,95	2,58	1,10	,571	
90- neutral report of the results	2,55	1,41	2,52	1,33	2,53	1,43	,791	
91- disclosure of results to all participants of the evaluation project	2,58	1,12	2,58	1,14	2,68	1,12	,403	
92- results are process-, not person-orientated	2,58	1,41	2,49	1,40	2,73	1,43	,133	
93- inclusion of all stakeholders when producing recommendations based on results	2,70	1,14	2,68	1,21	2,76	1,04	,283	
94- evaluator supports the usage of results	2,86	1,29	2,83	1,33	2,92	1,26	,572	
95- evaluation of the evaluation project itself in the sense of a meta-evaluation	2,86	1,22	2,81	1,21	2,99	1,23	,320	
96- results are client-friendly	2,88	1,58	2,79	1,61	2,97	1,54	,203	
97- report contains different scenarios how to use the results	2,91	1,16	2,78	1,20	3,08	1,08	,016	,017
98- results induce consequences	3,00	1,38	3,03	1,45	2,96	1,31	,864	

table 15 (continued)	all (442)		client (189)		eval. (232)		p	ω^2
	MR	SD	MR	SD	MR	SD		
99- client defines usage of results	3,09	1,38	2,89	1,32	3,31	1,42	,018	,016
100- results do not lead to negative consequences for the participants	3,31	1,71	3,26	1,76	3,37	1,64	,533	
101- report contains interpretations by all participants of the evaluation	3,32	1,35	3,20	1,33	3,43	1,34	,132	
102- report contains interpretations by the client	3,57	1,46	3,42	1,46	3,73	1,41	,070	
103- report contains no interpretation, only pure data analysis	5,17	1,25	5,11	1,34	5,29	1,05	,606	

Looking over all the tables one can summarize as follows:
To have a successful evaluation project it seems to be very important what (1) is evaluated and why (2), in which context (3) and with whom (4) an evaluation project takes place. It is important to clarify all this to have a good basis.
Furthermore it is important that the project is done professionally: One need a clear and adequate design (53), correct methodological procedures in the evaluation project (55) and a correct methodological procedure of data analysis (78). The project also has to consider ethics (54).
Of course the evaluator plays an important role. Incorruptibility (29), credibility (30), high evaluation-expertise (31) as well as objectivity and neutrality (33) of the evaluator are necessary.
A cooperative climate (21, 32) seems to be helpful, too.
Finally the results of the project are important of course, too. Results have to be complete and fair (76), the report has to describe the object of the evaluation as well as its context, objectives, procedure and results (77), results have to be reported in a comprehensible, receiver adequate way (79), the handling of results has to be sensitive, confidential, and factual (80), and results should be helpful and constructive (81).
This is not a complete list. Many other conditions are ranked only a bit worse and should be considered in practical work, too.
Comparing evaluators and clients one can see that both groups judge in a very similar way. There are some differences in detail, but there are too many statistical tests and to low effect size to give these differences a high importance.
Evaluators and clients seem to have a very similar view of successful evaluation projects.

RESULTS – EVALUATION PRACTICE

The second important interest of this study was to examine how the current evaluation practice is for all these conditions. As mentioned above the experts had 3 possibilities to rank the conditions (1 = receives too much consideration; 2 = receives appropriate consideration; 3 = receives too little consideration).
A first step to analyze the current evaluation practice is to look where the answers of the experts indicate a need for changes.

Again comparing the three groups (all, 442; client, 189; evaluator, 232) the following table shows all conditions where

- the mean rank is smaller than 3,0 (condition seems to be at least quite important)

and

- 50% and more of the experts say that this condition receives too little consideration (seems that some improvements should be made; 50% and more experts did not say anywhere that a condition receives too much consideration).

That means that we have a look at conditions now which appear to be quite important with a need of changing the current practice (see table 16).

Table 16:
Need for changes

	all (442)		client (189)		eval. (232)		p	ω2
	MR	% need	MR	% need	MR	% need		
establishment of an evaluation culture	2,75	57,7	2,72	62,3	2,84	53,3	,097	
usage of evaluation results is clarified in advance	2,38	53,8	2,42	52,5	2,40	55,2	,627	
clear and realistic evaluation objectives and questions	1,51	51,2	1,49	46,7	1,54	55,8	,090	
clear definition of the evaluation context	1,81	50,7	1,82	48,5	1,83	52,5	,458	

It casts a positive light on current evaluation practice that only four (quite) important conditions should be considered more in the future. These four indicate that the whole climate in which projects are done should be improved, and that some more detailed work regarding the usage of results and regarding the definition of objectives and of the evaluation context should be done before an evaluation project starts.
Again clients and evaluators are judging in a similar way.

A second step to analyze the current evaluation practice is to look where the answers of the experts indicate good practice.
Again comparing the three groups (all, 442; client, 189; evaluator, 232) the following table shows all conditions where

- the mean rank is smaller than 3,0 (condition seems to be at least quite important)

and

- 70% and more of the experts say that this condition receives appropriate consideration (seems that no improvements need to be made).

That means that we have a look at conditions now which appear to be quite important with no need of changing the current practice (see table 17).

Table 17:
Good practice

	all (442)		client (189)		eval. (232)		p	ω2
	MR	% ok	MR	% ok	MR	% ok		
sensitive, confidential, factual handling of results	1,89	80,7	1,77	75,0	2,00	86,7	,017	,018
the participants are accepted by the persons responsible for the project	2,78	79,9	2,87	79,3	2,77	81,5	,646	
data is accessable with justifiable effort	2,23	76,5	2,16	76,9	2,31	76,0	,856	
professional competence, familiarity of participants with the evaluation object	2,95	76,1	2,82	72,8	3,06	79,3	,213	
results are complete and fair	1,71	76,1	1,68	72,3	1,74	79,5	,170	
acceptance and credibility of the evaluator	1,78	75,4	1,72	73,6	1,85	75,5	,714	
responsibility of realization belongs to evaluator	2,91	74,3	2,97	69,1	2,82	80,0	,049	,012
report contains interpretations by the evaluator	2,51	74,0	2,44	73,1	2,58	74,6	,781	
comprehensible, receiver adequate report of results	1,89	73,4	1,80	70,0	1,97	76,4	,249	
voluntary participation of the participants	2,62	73,0	2,61	73,2	2,68	74,3	,842	
neutral report of the results	2,55	71,4	2,52	69,2	2,53	73,6	,432	
attention to data protection	2,14	71,2	2,12	68,4	2,18	73,3	,376	
report contains no rash conclusions	2,16	71,2	2,14	69,8	2,22	72,2	,666	
evaluator supports the usage of results	2,86	71,2	2,83	69,5	2,92	73,4	,498	

Looking at the conditions which show “good practice” one can see that nearly all parts of an evaluation project are included, with an emphasis on evaluation results where many details seem to work in practice. Taking into account that the cut off of 70% is chosen quite randomly and that many additional conditions can be found in a range of 60-70%, one can conclude that with the overall-look of this study evaluation practice seems to be on the right track. Again clients and evaluators are judging in a similar way.

SUMMARY

With the findings of this study first empirical indicators have been found for the question what an evaluation project needs to be successful.
To have a successful evaluation project it seems to be very important (among others) what is evaluated and why, in which context and with whom an evaluation project takes place. It is important to clarify all this to have a good basis.
Furthermore it is important that the project is done professionally: One need a clear and adequate design, correct methodological procedures in the evaluation project and a correct methodological procedure of data analysis. The project also has to consider ethics.
Of course the evaluator plays an important role. Incorruptibility, credibility, high evaluation-expertise as well as objectivity and neutrality of the evaluator are necessary.

A cooperative climate seems to be helpful, too.
Finally the results of the project are important of course, too. Results have to be complete and fair, the report has to describe the object of the evaluation as well as its context, objectives, procedure and results, results have to be reported in a comprehensible, receiver adequate way, the handling of results has to be sensitive, confidential, and factual, and results should be helpful and constructive.
This is not a complete list. Many other conditions are ranked only a bit worse and should be considered in practical work, too.

Comparing evaluators and clients one can see that both groups judge in a very similar way. There are some differences in detail, but there are too many statistical tests and to low effect sizes to name these differences as important. Evaluators and clients seem to have a very similar view of successful evaluation projects.

Regarding current evaluation practice only few, but important indicators for improvement have been found: Evaluation climate in which projects are done should be improved, and more emphasis should be put on the usage of results, on the definition of objectives and on the definition of the evaluation context before an evaluation project starts.

One can go in different directions now:

- Much additional statistical work can be done with the data. It could be interesting to take evaluation expertise into account in a sense to compare e.g. more practical orientated experts with more theoretical orientated ones. It could also be interesting to compare experts from different domains. But all this is too much for one paper.
- One has to have in mind that the questionnaire of this study has an over-all focus. It examines conditions over all domains, over all situations, over all evaluation objects. This is useful to get an overview and to get a feeling for mainstreams, but of course evaluation takes places in different situations with different characteristics so with data of this study it is not possible to take all possibilities of different situations and contexts into consideration. To analyse one specific domain or field new data could be helpful. Some participants of the study said that they have answered over-all, but they also have examples of evaluation-projects in mind where they would have judged in a different way if it would have been the task to answer the questionnaire for this specific context.
- Unfortunately it was not possible within this study to find enough experts who “only” participate in an evaluation project. Therefore, specific statements have only been possible regarding evaluators and clients. Other stakeholder groups should also be examined in detail, but here, more research is needed.

REFERENCES

Balzer, L. (2004). "Wie werden Evaluationsprojekte erfolgreich?" - Erste Ergebnisse einer Delphi-Studie. In M. Wosnitza, A. Frey & R. S. Jäger (Hrsg.), *Lernprozess, Lernumgebung und Lerndiagnostik - Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert* (S. 392-412). Landau: Verlag Empirische Pädagogik.