The Big Bad Wolf's View: The Evaluation Clients' Perspectives on Independence of Evaluations

Authors:
Lyn Pleger and Susanne Hadorn

Abstract:
The independence of evaluations and pressure put on evaluators by clients in particular have gained increasing attention in research. The call by the evidence-based policy movement (EBP) for the use of unbiased evidence within policy making highlights the importance of independent evaluations or, more specifically, of evaluations that are conducted in the absence of distortion. So far, research has mainly focused on evaluators' experiences when confronted with pressure, leading to a restricted view of the complex context in which evaluations take place. Therefore, this paper broadens the debate and follows the call by previous studies to pay attention to evaluation clients, which have been identified as the main influencing stakeholders within evaluation processes. The findings of an online survey among Swiss evaluation clients show that while clients agree that independence and evaluations standards are decisive for high quality results, most respondents have only limited knowledge about the standards' content. Surprisingly, results also show that most clients have never been told by evaluators that they put them under pressure, even though previous studies found that half of the evaluators feel pressured. Nonetheless, clients frequently experience conflicts during evaluation processes that are often caused by a lack of a common understanding between evaluators and clients. Thus, we conclude that preventive measures are needed to increase the parties' mutual comprehension, and ultimately, the quality of evaluation results. Only then can evaluations meet the principles of EBP.

1. Introduction

Independence of evaluations and potential threats to that independence have increasingly attracted attention in research over the last decades. The research strand of evaluation independence has its origin in ethical challenges during evaluation processes, as pressure on evaluators has often been identified as one important type of ethical challenges (Morris, 1999, 2007a, 2007b). Especially during recent years, many empirical studies examined the evaluators' perception of their independence (Morris & Clark, 2013; see also Pleger, Sager, Morris, Meyer, & Stockmann, 2016) and investigated the main source of pressure experienced by evaluators. One of the common findings of these studies was that the client is often the most influencing actor (Morris, 2007a; Morris & Clark, 2013; Stockmann, Meyer, & Schenke, 2011). Surprisingly, despite the growing research attention in the area of independence of evaluations and the consistent findings of clients playing a key role by putting pressure on evaluators, only very little attention has been paid on the client's perspective so far. Therefore, this gap is our starting point and raises the following research questions, which we attempt to answer in this paper: Are clients aware of the pressure they are putting on evaluators and how do they perceive their relationships? How do clients assess the importance of evaluation independence and how familiar are they with evaluation standards? Moreover, what kind of measures do clients propose to create a more fruitful environment for meaningful evaluations and to what extend do they agree with preventive actions suggested by evaluators? These questions are of particular importance, because evaluations can only become more independent, if both sides – evaluators and clients – are aware of the challenges they face in their relationships. Due to a lack of research in the area of client perception of independence and evaluations, the present study provides the first descriptive insight into client perspectives and a comparison between the experiences of evaluators and clients. In doing so, it is not the aim of this pilot study to identify causalities and empirical relationships but to lay foundations for future research efforts by presenting evidence from an online survey of Swiss evaluation clients. The paper starts with a

discussion of the state of research of pressure on evaluators as an ethical challenge, followed by the presentation of the research design. After presenting the data collected, the results are discussed. The article closes with a presentation of the main conclusions and directions for future research.

2. Evaluations and evidence-based policy making

Evaluations ideally seek to make an unbiased and independent assessment of the underlying subject matter, such as public policies, to provide decision makers with the best possible factual basis for policy decisions. This idea of gathering objective evidence through evaluations is inherently connected with the notion of evidence based policy making (EBP) that searches for the "continuous improvement in policy settings and program performance, on the basis of rational evaluation and well-informed debate of options" (Head, 2008, p. 1). However, Sanderson (2002, p. 5) asserts that, "the ideal model of evidence-based policy making is predicated upon certain assumptions relating to [...] the ways in which evaluation can provide the evidence needed". Thus, one part of this assumption is that evaluators are capable – and are given the freedom - to gather objective evidence that help to optimize policies in the future, or in other words that "evaluation results are politically unbiased" (Pleger & Sager, 2016c, p. 1). Equally, however, it seems obvious that evaluations take place within a political context, as they are usually commissioned by third parties such as politicians or public servants, and are not self-initiating processes. In light of this, evaluation literature has a long tradition of studying ethical challenges faced by evaluators that might ultimately hamper the quality of evidence collected (see e.g. Brown & Newman, 1992; Morris, 1999; Morris & Jacobs, 2000; Newman & Brown, 1996; Pope & Vetter, 1992; Stufflebeam, 1994). A study by Morris and Cohn (1993) including a questionnaire among American evaluators found that conflicts regarding the reporting of evaluation results were the most frequently mentioned ethical problems. Specifically, such problems "involve attempts on the part of a client/key stakeholder (as

perceived by the evaluator) to undermine the fundamental mission of scientific inquiry, which is to seek the truth and communicate it" (Morris & Cohn, 1993, p. 639). Similarly, Turner (2003) stated that issues such as control or influence of evaluation results by the commissioning party were among the most prevalent ethical challenges faced by Australasian Evaluation Society members. In the search for a closer understanding of dynamics within contracted evaluations, recent studies have thus increasingly focused on this specific ethical challenge, the independence of evaluations. Within this paper, we will focus our analysis on this aspect.

3. State of research: pressure as the threat of independent evaluations

Studies investigating ethics of evaluations primarily focused on the point of view of evaluators. A notable exception is the study by Morris (2007b) which has investigated foundation officers' (evaluation clients') experiences with ethical challenges during evaluations. The findings showed that ethical challenges faced by the respondents included, amongst others, disagreements about the evaluation results, the dissemination and ownership of evaluation findings and issues regarding the evaluators' motives and their work quality. In regards to the presentation of findings, Morris (2007b) found that while previous studies showed that evaluators have often the feeling that they are put under pressure to misrepresent findings, the results of the survey of evaluation clients suggest another cause for misrepresentation. Specifically, evaluators themselves might "engage in the same sort of misrepresentation of results that they so often accuse other stakeholder of encouraging them to undertake", which is why we should not depend "solely upon the views of evaluators when mapping the terrain of ethical challenges in the field" (Morris, 2007b, p. 413). Despite this implicit call by Morris (2007b) to pay more attention to clients' perspectives on issues of (mis)representation and thus independence of evaluations, in the recent body of literature about the independence of evaluation, researchers were so far primarily interested in the views of evaluators. Specifically, similar studies were conducted within the USA, the United Kingdom, Germany and

Switzerland to examine questions such as how evaluators experience pressure from third parties and how they react to it (Morris & Clark, 2013; Pleger & Sager, 2016a; Stockmann et al., 2011; The LSE GV314 Group, 2013). A comparative study including the data of all four aforementioned surveys found that, if asked in the surveys, responding evaluators showed strong commitment towards the evaluation standards and to the principle of always reporting the actual evaluation findings (Pleger et al., 2016). Also, a large proportion of German and Swiss evaluators disagree with the statements that evaluation results should be represented in a diluted form or adjusted to the expectations of clients (Pleger et al., 2016). Importantly, however, while this study suggests that professional ethics within the evaluation community is high, it also emphasizes potential threats to evaluation independence, specifically through pressure exerted by clients. The various empirical studies found that evaluators are often confronted with influences, whereby evaluation clients were identified as the most important factor of influence (Morris & Clark, 2013; Pleger & Sager, 2016a; Stockmann et al., 2011). In particular, 42, 50 and 78 percent of American, Swiss and German evaluators respectively indicated that they have experienced pressure to misrepresent findings at least once in their career, while most respondents replied to have experienced pressure on several occasions (Pleger et al., 2016). However, as Morris and Cohn (1993, pp. 637–638) pointed out, being pressured does not always result in ethical violations, "because some problems are handled in a way that prevent violations from occurring. Thus being pressured to distort evaluation findings or breach confidentiality is not same thing as succumbing to it". The critical question that has to follow is thus how evaluators handle pressure. Morris and Clark (2013) argue that simply resisting pressure requires, firstly, a commitment to ethical standards, secondly, the awareness of potential consequences when following these principles and, thirdly, the readiness to bear these consequences. While the first factor of high ethical standards is according to the aforementioned empirical studies given in most cases, the degree of risk preparedness or aversion might strongly depend on each evaluator and the given circumstances. In regards to

evaluators' resistance to pressure Pleger et al. (2016) concluded that the vast majority of evaluators stated they have made changes due to influences from stakeholders. However, while most of these adjustments did not (according to the respondents) constitute a way of misrepresenting results, 16 percent of American and 9 percent of Swiss evaluators admitted to already having misrepresented findings due to pressure from third parties.

In summary, dynamics between evaluators and clients are critical for the notion of EBP as "the attempt of a client to influence the evaluation process in terms of the data used, methods, results or conclusions can result in a distortion of the evidence provided by the evaluation" (Pleger & Sager, 2016c, p. 3). Given that recent studies have highlighted the frequency in which evaluators feel pressured, it is therefore crucial to also appreciate the other side of the coin: the clients' point of view. This seems to be decisive for two reasons: Firstly, a better understanding of their perspective might initiate a more insightful dialogue between evaluators and clients. Or as Morris (2015, p. 38) puts it: "[m]ore detailed information on the nature of […] expectations, beliefs, and attitudes [of clients] can equip evaluators to interact with stakeholders in ways that are more ethically proactive than is typically the case, especially during the crucial entry/contracting stage of the evaluation". Secondly, by initiating this discussion, current deficits in our understanding of client awareness of unethical behavior might ultimately lead to higher levels of evaluation independence.

4. Method

The data presented in this paper was collected in an online-survey that was carried out between March and April 2017 using the software *Qualtrics*. Since clients are less organized than evaluators (e.g. in evaluation societies), the e-mail invitation was sent via different channels including a non-individualized link. The survey invitation was sent to all cantonal offices and evaluation experts in national offices as well as to private sector companies and non-profit organizations which had commissioned evaluations in the past. The invitation contained a

request to further distribute the link to colleagues which have commissioned evaluations in the past. The means of survey distribution, however, does not allow any conclusions regarding the response rate.

By changing the perspective from evaluators to clients, most of the questions in the survey were the counterparts of those in the questionnaire of the study by Pleger and Sager (2016a), which investigated the independence of evaluations from the evaluators' perspective. The similarity of the questions allows for a comparison between the perception of evaluation independence of evaluators and clients. Due to the multilingualism in Switzerland, the e-mail invitation as well as the survey were provided in German and French. The survey consisted of 43 questions assigned to five thematic blocks addressing work experience, ethical values within evaluation practices, pressure on evaluators, familiarity with evaluation standards and evidence-based policy making and a final block pertaining to socio-demographical characteristics.

Respondents were told in the invitation that the survey addresses "independence of evaluations". Within the survey, the meaning of the term "influence" in the block covering pressure and influence on evaluators and independence of evaluations, was defined as referring to the negative connotation only.[1]

In total, 152 individuals participated in the survey. However, it was stated in the e-mail invitation that the survey only targets clients of evaluations and not evaluators. The beginning of the survey contained filter questions for identifying whether the individual commissioned evaluations or not. All respondents who stated to either work as an evaluator rather than commissioning evaluations or who did not have any work experience were excluded from the sample. The final dataset consisted of 56 respondents. The main reason for the limited number of respondents lies in the fact that in contrast to evaluators, clients are much less organized. While evaluators are widely organized within the Swiss evaluation society (SEVAL), there is

---

[1] The wording was the specified influence as follows: „The concept of "*influencing*" can have a positive as well as a negative meaning for many people. For this reason, we would like to point out that our questions always aim at *negative* influences and influence is used interchangeable with the term "pressure" throughout this survey."

no comparable counterpart for clients. Consequently, it is much more difficult to reach a large number of them.

5. Results

With the exception of the filter questions, the questions in the survey were not mandatory, thus the sample sizes vary depending on the question. The results presented are based on valid responses only. Hence, non-respondents were removed from the data.

5.1 Respondents characteristics and employment setting

Regarding the socio-demographic characteristics, it is worth noting that male respondents were overrepresented with 76 percent (n = 29) compared to female respondents with 24 percent (n = 9). Clients were relatively highly educated as 47 percent held a university degree, and 32 percent a Ph.D. Only one person (3 percent) stated to not have a degree, 16 percent held a degree from a university of applied science and one person (3 percent) held a habilitation (N = 38).

Overall, the respondents had a relatively long professional experience according to their number of years commissioning evaluations: The mean number of years commissioning evaluations was 9.4 (*S.D.* = 7.1; N = 53). The employment setting of the respondents was rather homogeneous as 92 percent stated to work in the public sector for the federal, state or local agency (except colleges or universities), 6 percent worked for a nonprofit organization (NGO) and 2 percent claimed to be self-employed (N = 52).

The majority of the respondents (88 percent, n = 45) commissioned mainly external evaluations in comparison to 12 percent (n = 6), who mainly commissioned internal evaluations. More specifically, the main contractors worked in the private sector, followed by universities and colleges or were self-employed. Further listed contractors were the public sector except universities and colleges and NGOs. The option "other" included the open-ended responses "evaluation offices" and "schools" (Figure 1).

Figure 1 – Primary contractor of respondents
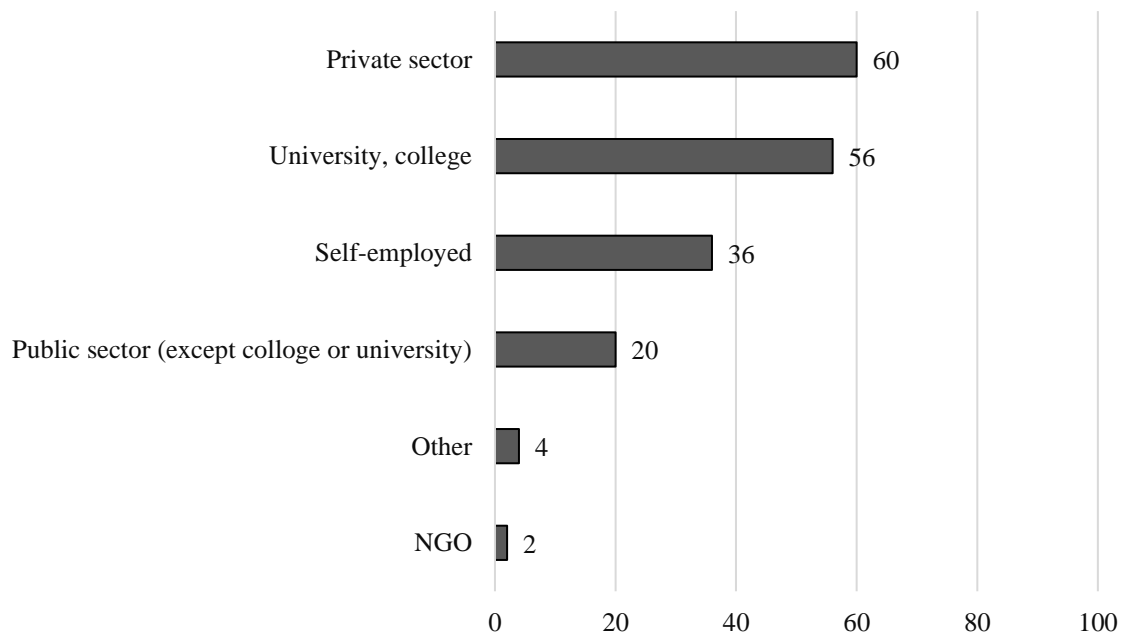
Notes: Numbers represent percentages (N = 50); multiple answers were possible.

Clients were also asked about the procedures they apply to tender evaluations. The top award procedure was Tenders (82 percent) followed by Direct awards (70 percent). 48 percent of the respondents awarded contracts via Public tenders, 26 percent via WTO-tenders and 2 percent stated "Other" (Figure 2).

Figure 2 – Award procedures for evaluations
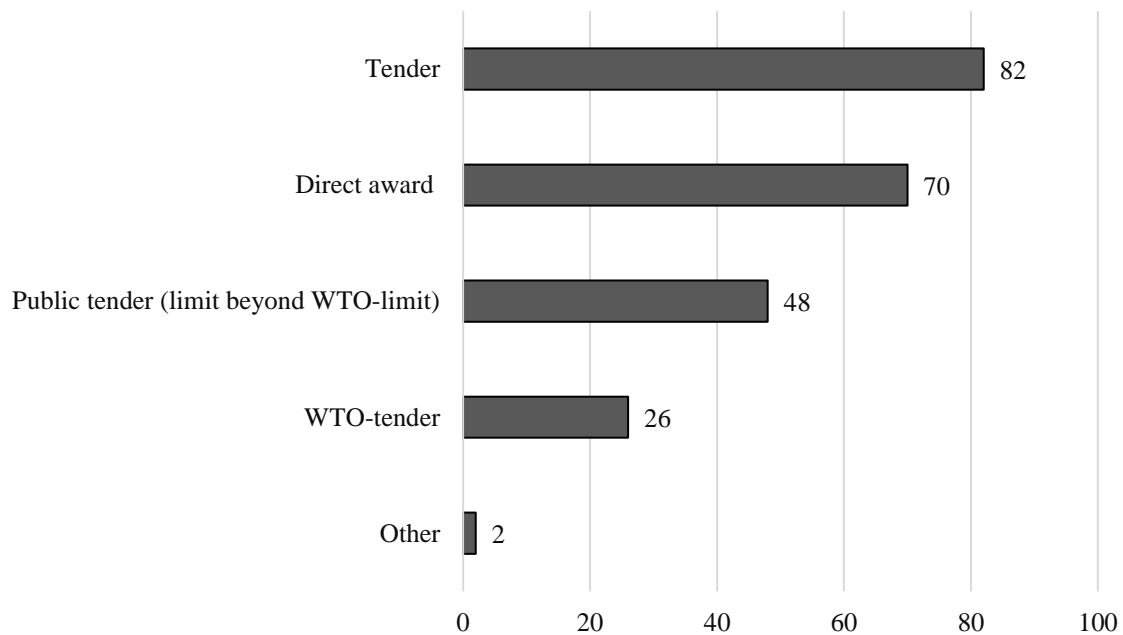
Notes: Numbers represent percentages (N = 50); multiple answers were possible

An additional question asked specifically for the percentage of evaluations that are awarded directly. Accordingly, of those indicating that they award contracts directly, 53 percent (n = 18) awarded up to half of their evaluation contracts directly while 47 percent (n = 16) of the respondents awarded between 75 percent and 100 percent of the evaluation contracts directly.

5.2 Guiding principles and ethical standards

In order to investigate the perception of clients regarding independence of evaluations in general and their pressure on evaluators in detail, it is important to gain insights into the clients' assessments of ethical standards for evaluators. In order to compare the clients' perspectives with those of evaluators, the clients were asked similar questions posed to evaluators in the study by Pleger and Sager (2016a) to indicate their degree of agreement to different ethical standards for evaluators. Table 1 shows client and evaluator mean approval of different statements addressing ethical standards for evaluators.

Table 1 – Perception of ethical standards for an evaluator

| A good evaluator…	Clients			Evaluators		
	M	*S.D.*	*N*	*M*	*S.D.*	*N*
……always only presents the actual results, without diminishing or enhancing them	5.3	0.9	42	5.0	1.2	102
...is primarily obliged by the evaluation standards.	5.0	1.1	42	5.3	0.8	103
… relentlessly reveals determined weaknesses in the evaluation	5.0	1.2	44	4.7	1.2	102
... does not respond to influence regarding the methodical approach	4.6	1.1	41	4.8	1.1	102
…adjusts evaluation results according to the expectations and needs of the clients	3.6	1.7	43	2.9	1.6	98
… primarily has a moral responsibility to the stakeholders	2.9	1.5	39	3.3	1.2	94

Notes: Numbers are based on a scale between 1 (= strongly disagree) and 6 (= strongly agree); Data for evaluators taken from Pleger and Sager (2016a, p. 32)

The highest degree of client approval received was the statement that "a good evaluator always only presents the actual results, without diminishing or enhancing them", whereas evaluators gave a higher approval to the statement that "a good evaluator is primarily obliged by the evaluation standards". However, these two statements found a high mean approval by both clients and evaluators. Interestingly, a statement which received a distinctly higher mean approval by the clients than evaluators was that "a good evaluator adjusts evaluation results according to the expectations and needs of the clients". Conversely, the statement that "a good evaluator primarily has a moral responsibility to the stakeholders" was more strongly approved by the evaluators compared to the clients. Very similar were the results for clients and evaluators for the approval of the statement "a good evaluator does not respond to influence regarding the methodical approach" (Table 1).

5.3 Independence of evaluations

In order to tackle independence of evaluations concerning pressure on evaluators, clients were firstly asked for the main difficulties when collaborating with evaluators. Clients' answers to this open-ended question were very fruitful. In total, 30 clients responded to that questions, some at length, thus the responses provide a comprehensive insight to the perception of difficulties from clients' perspectives in working with evaluators. Some of the responses included several aspects, thus, in total 36 difficulties were described. These were assigned to five different categories and are shown in Figure 3.[2] More precisely, 42 percent of the open-ended responses could be assigned to the reasons referring to a lack of evaluators' understanding of the organization that was evaluated or a lack of a mutual understanding between the evaluator and the client (Figure 3). Examples for responses in this category are that evaluators would have a "partly lacking understanding of the organization to be evaluated (regarding processes etc.)" or evaluators are deemed to have "lacking / insufficient detailed knowledge about the subject of the evaluation."

Figure 3 – Perceived main difficulties when working with evaluators
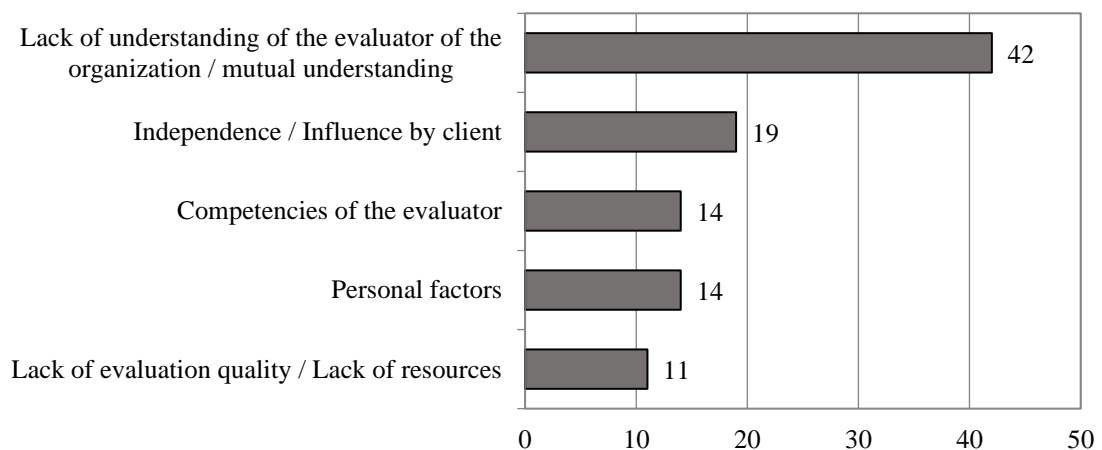
Notes: Numbers represent percentage; Open-ended responses were assigned to one of five categories (N = 36).

---

[2] The reasons were coded and assigned to the categories by both authors independently. The coding revealed an inter-coder reliability (percent agreement) of 92 percent (r_c = 33/36 = .92), which corresponds to a very high convergence coefficient (Neuendorf, 2016, p. 181; Raupp and Vogelgesang, (2009, xiv).

Interestingly, the category with the second most frequently given open-ended responses concerns the "independence of evaluations and influence of the client", which comprises 19 percent of the responses. For instance, one client responded to the question for the main difficulties in working with evaluators with "letting the evaluator do his job" or "wishes regarding the content and results of the evaluation". Another client stated the main difficulties when working with evaluators were the "tension between *providing services* and *ensuring independence*". More generally, another respondent said difficulties when working with evaluators include the fact that "evaluators can be manipulated in their judgment, whether it be in the course of the evaluation by other persons involved or by pre-existing prejudices of the evaluator that have not been made transparent".

14 percent of the open-ended responses addressed difficulties relating to evaluator competencies such as a "lack of professional competencies of the evaluator". Another client stated "a lack of transparence of the evaluator regarding theoretical, methodological or interpretational aspects" as an example for difficulties.

An example for a response given that was assigned to the category "personal factors" which contains 14 percent of the open-ended responses were difficulties to prevent "that evaluators do not incorporate their political views and personal preferences into the evaluation into an evaluation". The fifth category "lack of evaluation quality / lack of resources" contains 11 percent of the responses, for example "time pressure of the evaluators". One respondent argued that "evaluators have too many assignments and therefore are under constant stress".

When asked about putting pressure on evaluators in the context of evaluations, the results from the survey on clients differ greatly from the survey on evaluators (see Figure 2, Pleger & Sager, 2016a). While almost none of the clients (92 percent) have ever been accused by an evaluator that he or she put pressure on the evaluator or influenced him or her to display results incorrectly or inaccurately, 50 percent of the evaluators indicate that they have experienced such pressure.

Figure 4 – [a] Clients who were accused of putting pressure and [b] evaluators experienced pressure

[a] Clients
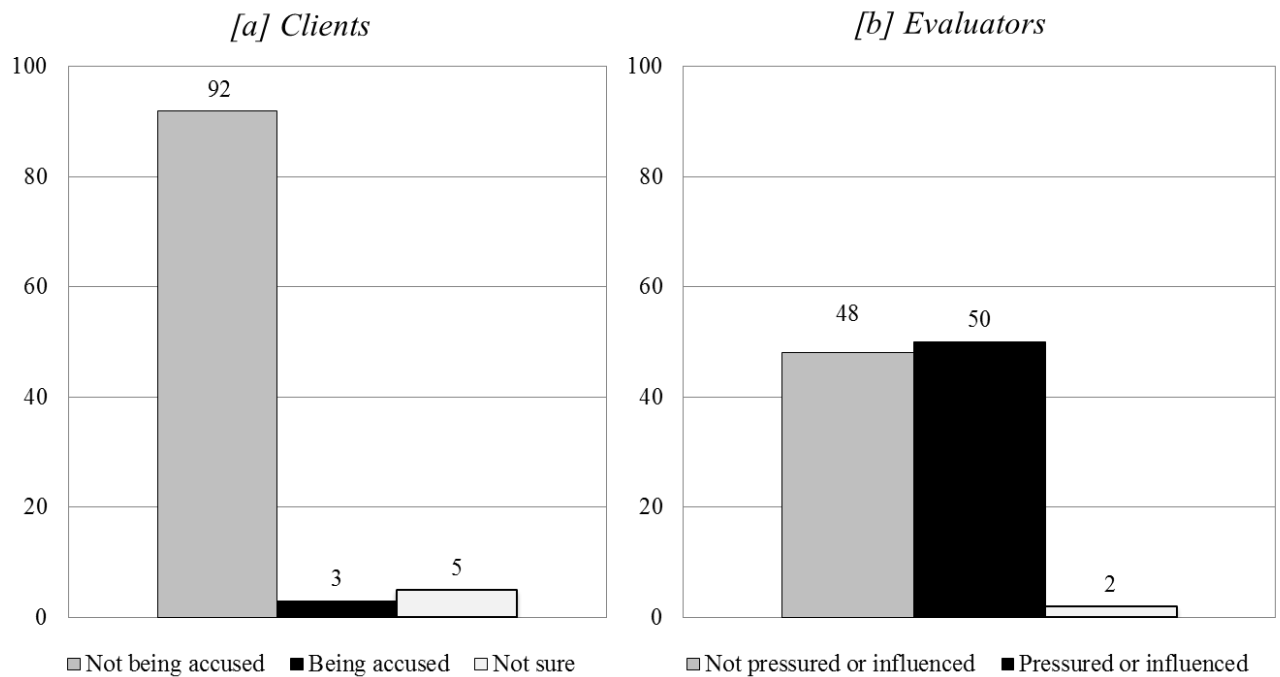
[b] Evaluators

Notes: Results represent percentages; results for evaluators taken from Pleger and Sager (2016a, p. 33); (Clients N = 39; Evaluators N = 123)

In a next step, clients were asked if they have ever been dissatisfied with a commissioned evaluation and if they ever had a conflict with evaluators due to the latter's work. Among respondents, 53 percent (n = 18) reported that they have been dissatisfied with an evaluation they have commissioned whereas 47 percent of the clients have never been dissatisfied with an evaluation (n = 20). Regarding conflicts, 56 percent (n = 22) never had a conflict with an evaluator and 44 percent (n = 17) had. As shown in Figure 5, the results for these questions are very similar which might indicate that clients tend to communicate their dissatisfaction to the evaluators which, in turn, leads to conflicts.

Figure 5 – Dissatisfaction with evaluations (N = 38) and conflicts with evaluators (N = 39)

Have you ever been dissatisfied with an evaluation commissioned by you?

Have you ever had a conflict with evaluators due to their work?
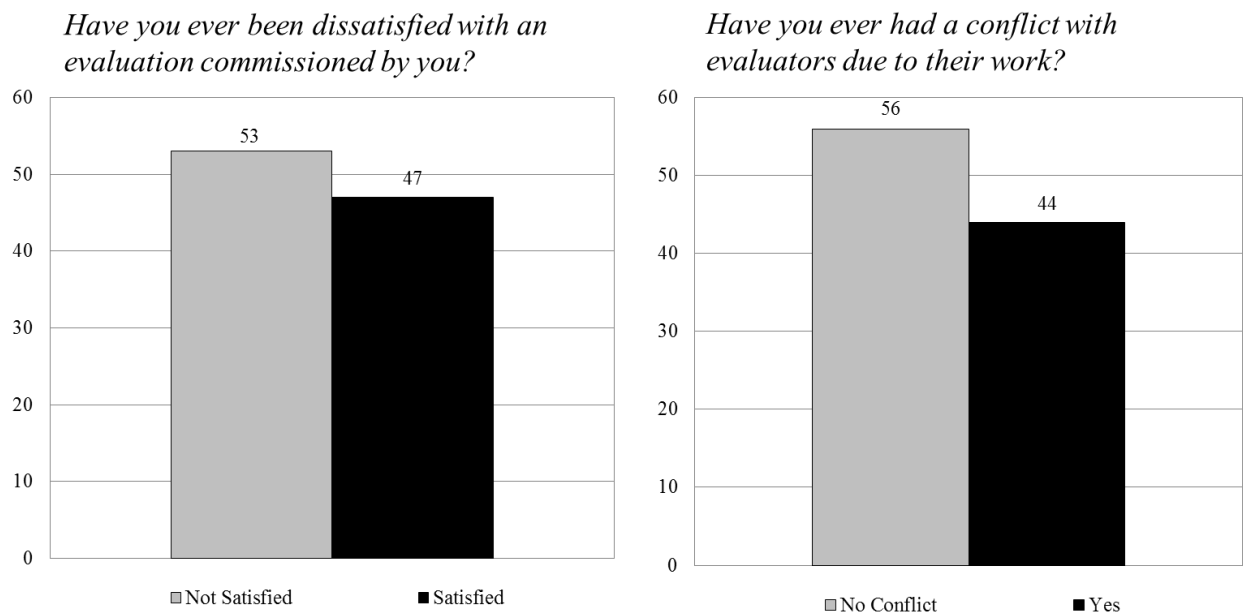
Regarding the cause for the dissatisfaction with an evaluation, 63 percent claimed that quality of the evaluation was insufficient, 16 percent stated that the evaluation did not meet expectations and 5 percent said the schedule was not being adhered to. Moreover, 16 percent gave other reasons as open-ended responses (N = 19) such as "The evaluation execution was insufficient" or „The evaluator did not understand the context".

Those clients who have had a conflict with evaluators were also asked for the reason for this specific conflict. The responses to this open-ended question provided a broad variety of causes of conflicts. The range of conflict causes encompasses a "lack of (methodological) competences of the evaluator", "a lack of quality of delivered results" and "Disagreement with essential parts of the evaluation results". Further examples for given causes for conflicts were: "The report contained unacceptable political assessments", "Differing views on the political mandate and the framework conditions" and "No understanding of the requirements of the client".

To understand the different forms of influence, respondents were also asked in closed questions to indicate on a scale between 0 (= No, I've never done that) and 3 (= Yes, I've often done that) if they have ever proposed different amendments. Table 2 contains the amendments clients

proposed on at least one occasion. Around one third (31 percent) stated that they have suggested to present evaluation results more positively in one or more than one occasion. Most of the clients, 97 percent, suggested at least once that individual evaluative sentences should be reworded in the evaluation report in at least one occasion. Further, 55 percent suggested that other conclusions should be drawn from the results in at least one occasion. Additional data sources were suggested by 80 percent on at least one occasion, while 25 percent suggested changes should be made that would lead to content distortions on at least one occasion, of which 3 people stated "Yes, I've often done that". Finally, 97 percent stated that they have never informed the evaluator in advance of the evaluation of the results they expect, while one person stated to have done it often.

Table 2 – Proposed amendments by the clients

	Percent	Frequency	N
Rewording of individual evaluative sentences in the evaluation report	97	30	31
Use of additional data sources	80	24	30
Drawing of different conclusions from the results	55	17	31
Present evaluation results more positively on one or more occasions	31	10	32
Changes would lead to content distortions on at least one occasion	26	8	31
Informed the evaluator in advance of the evaluation of the results they expect	3	1	32

Notes: Proposed amendments by clients on at least one occasion; multiple answers were possible

Additionally, clients were asked in an open-ended question for further amendments they have proposed. For example, one client stated that he or she asked the evaluator to "remove political and evaluative statements from the report in order to preserve objectivity". Another client said he or she instructed the evaluator to "pay attention to political sensitivities in the formulation of results in order to achieve greatest possible acceptance". Another example of a given proposed amendment was that the client said he or she requested the "rectification of interpretations".

Respondents were also asked how evaluators' reacted to these modification requests. In the majority of instances (53 percent), the client and the evaluator found a compromise, 44 percent stated that the changes were done by the evaluator and in 3 percent of instances, the changes were not done (N = 32).

5.4 Evidence-based policy making, evaluation and evaluation standards

As discussed at the beginning of this paper, this study seeks to contribute to research on independence of evaluations in the context of evidence-based policy making. Hence, clients were asked whether they were familiar with EBP. Surprisingly, the majority of clients (69 percent) stated to not be familiar with the concept of EBP compared to 31 percent who were familiar with EBP (N = 39). Those who were familiar with EBP were then asked to consider the importance of complying with EBP within their area of work. Ultimately, all clients except one indicated that their compliance with EBP is reasonably important or very important for their work (N = 12).

In evaluation practice, however, the underlying notion and ideals of EBP are expressed by evaluation standards developed and published by evaluation societies. Thus, clients were also asked about their familiarity with the Evaluation Standards by the Swiss Evaluation Society SEVAL (SEVAL standards, 2017). Over two thirds of the clients (68 percent) reported to know of the SEVAL, while 32 percent did not. Those who knew of the SEVAL Evaluation Standards were asked how familiar they were with them. Accordingly, 37 percent stated to be not familiar at all or rather partially familiar with the evaluation standards and 63 percent were reasonably familiar or very familiar with the evaluation standards. When respondents were asked how important the evaluation standards were for the evaluation process and the evaluation result, none of the clients stated that the standards were not important at all, 15 percent assessed the

evaluation standards as reasonably unimportant, 46 percent as rather important and 39 percent as very important (N = 26).

5.5 Preventive actions

Preventive actions that help to reduce or inhibit damaging influence on results can be decisive for the independence of evaluations. The closed question about potential preventive measures was identical to the one in the Swiss survey of evaluators (Pleger & Sager, 2016a), to allow for a direct comparison between the evaluators' and the clients' perspectives. The results are shown in Figure 6.

Figure 6 – Approval of potential preventive action
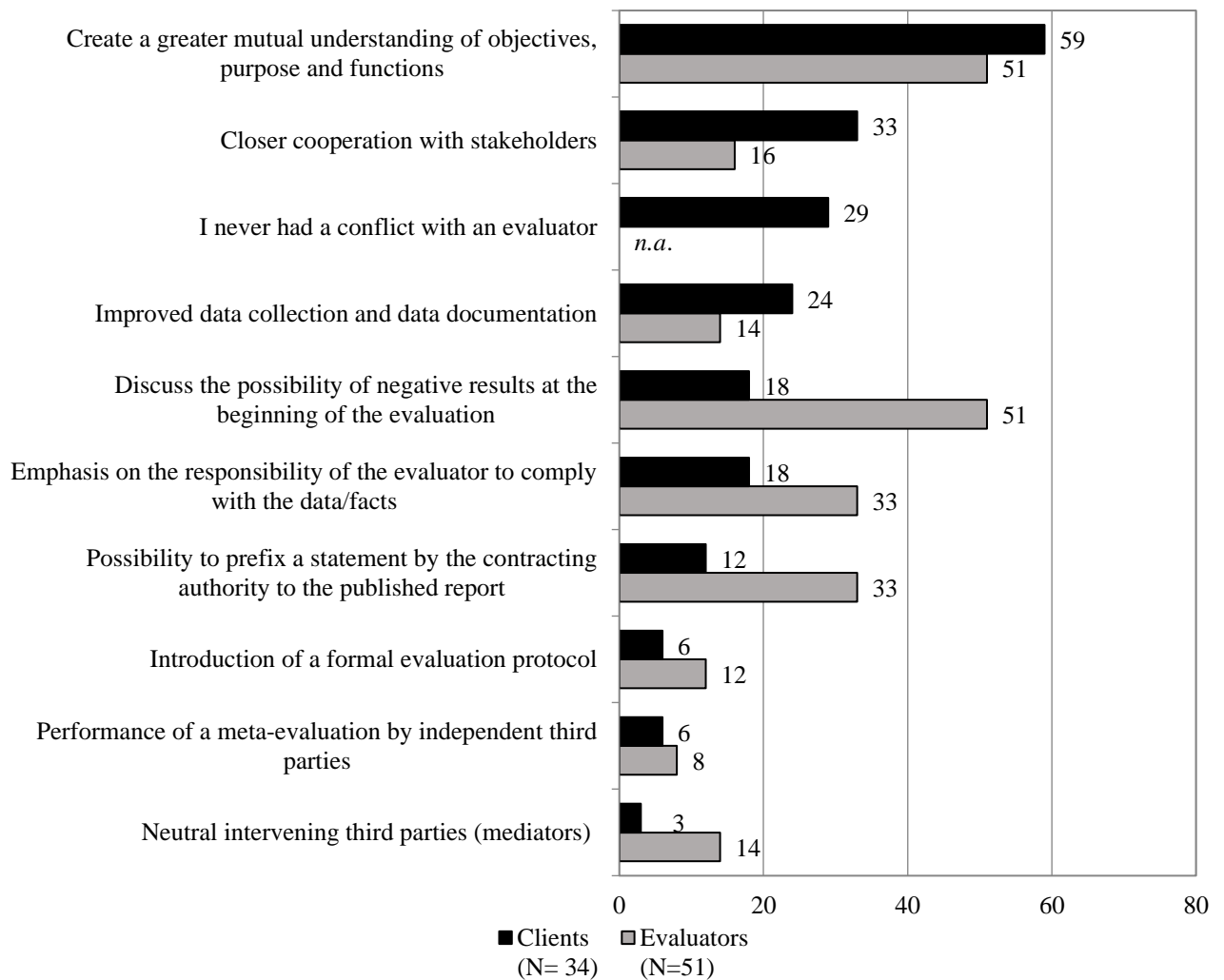
*Note*s: Numbers represent percentage; multiple answers were possible; Data for evaluators taken from Pleger and Sager (2016a, p. 49)

Interestingly, clients and evaluators agreed on one preventive action, which found particularly high approval, namely to "create a greater mutual understanding of objectives, purpose and functions" (59 percent clients, 51 percent evaluators). In contrast, the preventive action to "discuss the possibility of negative results at the beginning of the evaluation" was approved by about half of the evaluators (51 percent) but only by approximately one fifth of the clients (18 percent). Further diverging approval rates between clients and evaluators appeared for the preventive action of a "closer cooperation with stakeholders", which was supported by one third of the clients (33 percent) but only by 16 percent of the evaluators. Conversely, the preventive actions of a "possibility to prefix a statement by the contracting authority to the published report" and an "emphasis on the responsibility of the evaluator to comply with the data/facts" was approved by one third (33 percent) of the evaluators for both statements, but only by 16 percent and 18 percent respectively by the clients. The lowest degree of approval by clients received the preventive action of "neutral intervening third parties (mediators)". Only 3 percent of the clients considered such mediators as a suitable means to prevent conflicts with evaluators, whereas 14 percent of evaluators approved mediators as a preventive action (Pleger & Sager, 2016a). Moreover, 9 percent proposed other potential preventive actions.

6. Discussion

The responses from the survey reveal several interesting findings concerning both the evaluation procurement procedure and the actual evaluation process. Firstly, one aspect regarding the tendering process of evaluation contracts appears to be particularly important for questions of independence. The survey answers show that 70 percent of clients award the majority of their evaluations through direct awards, meaning that the choice of the evaluator is strongly based on their own judgement. This large margin of discretion is crucial because clients also stated that previous experiences with evaluators are decisive for the selection procedure. Bearing in mind that evaluators are financially dependent on evaluation contracts, a practice of

direct awards thus might increase the likelihood of giving in to pressure. As Morris and Clark (2013) argue, the decision to whether resist to pressure or not is not only the results of the evaluators' commitment to ethical standards, but also based their preparedness to take the associated risk. At the same time, the risk of losing future contracts due to prior conflicts is higher in cases where these are assigned by means of direct awards. Hence, knowing that the future order situation and business success strongly depends on the level of satisfaction of clients with the current project potentially motivates evaluators to give in more often to pressure from the clients.

Turning the actual process of evaluating and filing the respective report, the results reveal that the vast majority of clients has already asked for certain amendments. Most respondents have proposed changes that concern rewording of certain sentences, the use of different sources or reinterpretations within the reports. Importantly, approximately a quarter of the respondents already asked at least once for changes that would lead to content distortions. This number clearly emphasizes the relevance and particularly the prevalence of the issue of influence. However, what is striking in this is the remarkably small number of clients who indicated that they have already been confronted with the allegation of putting evaluators under pressure. Against the background that the review of Pleger et al. (2016) showed that many evaluators have already experienced pressure from clients, only a negligible share (3 percent) of clients indicated that they have at least once been confronted with the accusation of pressuring the evaluation team. This discrepancy between evaluators' and clients' responses can be explained in two ways: Either clients have chosen to give socially desirable answers or, which seems to be much more likely, evaluators only confront their clients very rarely with regard to this substantial issue. If the latter is true, effective conflict communication between clients and evaluators is a core deficiency of evaluation practice. A lack of information and communication between the two actor groups also becomes apparent in other areas, such as the perception of each other's roles. On the one hand, the vast majority of the clients agree with the statement

that a good evaluator is primarily obliged by the evaluation standards. On the other hand, even though clients recognize the importance of evaluation standards, 32 percent indicated that they do not know the standards and 37 percent of those who indicated knowing them have only a little knowledge of their actual content. It is therefore striking how severe the need for information about ethical standards still is on this side of the evaluation community.

Interestingly, the problem of insufficient communication seems to be recognized by both evaluators and clients. This is reflected in the fact that both groups indicate that the creation of a "greater mutual understanding of objectives, purpose and functions" is their preferred measure to ensure higher independence. On the contrary, one particular other preventive measure that could also contribute to an increased mutual understanding, the involvement of and mediation by a third independent party, is much less popular amongst both parties. This is of particular importance as it seems optimistic to hope for an ideal balance of power between evaluators and clients that would lead to an optimal *inclusion* of constructive insights and a simultaneous *exclusion* of misleading information into evaluation results. It is exactly this dilemma that seems to call for a mediating instance, for example in the form of a specialized evaluation service that operates at the interface between client and evaluator. One question remains therefore unanswered: Why, even though rationally justifiable, is the idea of assigning a mediating party for both clients and evaluators as unalluring?

Finally, this study has demonstrated the importance of the varying notions of pressure and influence. This survey showed that influences on the part of clients can clearly be negative if they constitute misrepresentation. However, feedback from clients can in other cases also beneficial to the quality of evaluation results, as misunderstandings or analytical errors on the part of evaluators can be corrected with useful insights. This goes in hand with the BUSD heuristic model by Pleger and Sager (2016c; see also Pleger & Sager, 2016b) that illustrates the diverse facets of influence. With a view to the survey results discussed in this paper, clients

often have the impression that evaluators do not fully understand the complexity of the evaluation target. In fact, this issue was the most frequently mentioned difficulty in their relationship with evaluators. It is therefore no surprise that clients, who fear that results might be invalid, try to achieve an improvement of the accuracy of evaluation results through implicit or explicit influence. In the discussion of independence of evaluations it is therefore crucial to distinguish between positive and negative forms of influence that ultimately determine whether alterations in evaluation results constitute misrepresentation or lead to an enhanced quality of findings.

7. Conclusion

The aim of this study was to contribute to one aspect in the field of independent evaluations that has been neglected by research so far: Whereas scholars have increasingly focused on evaluators' perceptions of being influenced during evaluation processes, the alleged source of such influence, the clients, has not been subjected to a more detailed examination. This is where this study contributes to by providing first enlightening insights. It is important to bear in mind, however, that the findings of this study are limited to some extent regarding their external validity. Due to the small sample size and a skewed sample concerning socio-demographic characteristics (76 percent male versus 24 female respondents) and the employment setting (mainly public sector), the sample is not representative of all Swiss clients. However, this paper used Switzerland as a case study because previous findings on evaluators' perspectives have shown that the Swiss case is closely comparable with other Western countries such as the US or Germany (Pleger et al., 2016). Despite the small sample size, this study has still produced valuable findings for the research field of independence of evaluations. Overall, findings revealed that from clients' perspective conflicts and difficulties appear when working with evaluators. Also, clients have stated that they clearly try to influence the evaluation process, although the underlying intentions can be of different nature. While some clients are

consciously striving for a misrepresentation of evaluation findings, in other cases it can be assumed that such influence is not necessarily intentional or even seeks for positive and beneficial changes. In this regard, the low level of knowledge of clients about the evaluation standards is striking, thus suggesting that clients are often not aware of the ethical problems they are causing through interference in the evaluation process. Future research should therefore investigate how underlying principles of EBP in general and evaluation standards in particular can be made more accessible for clients. Going a step further, this paper has clearly shown that insufficient communication between clients and evaluators is at the root of the problem; be it regarding a mutual understanding of the evaluation task or in the case of conflicts. The notion of the 'evil' client and the 'good' evaluator is thus too simplistic, even though action is clearly necessary to counteract negative dynamics between the evaluating and the evaluated parties. This study therefore strongly recommends that communication mechanisms between evaluators and clients are reconsidered and optimized to allow for more independent evaluations.

References

Brown, R. D., & Newman, D. L. (1992). Ethical Principles and Evaluation Standards: Do They Match? *Evaluation Review*, *16*(6), 650–663.

Head, B. W. (2008). Three Lenses of Evidence-Based Policy. *Australian Journal of Public Administration*, *67*(1), 1–11. https://doi.org/10.1111/j.1467-8500.2007.00564.x

Morris, M. (1999). Research on evaluation ethics: What have we learned and why is it important? *New Directions for Evaluation*, *1999*(82), 15–24.

Morris, M. (2007a). *Evaluation ethics for best practice: Cases and commentaries*: Guilford Press.

Morris, M. (2007b). Foundation officers, evaluation, and ethical problems: A pilot investigation. *Evaluation and Program Planning*, *30*(4), 410–415.

Morris, M. (2015). Research on evaluation ethics: reflections and an agenda. In P. R. Brandon (Ed.), *Research on evaluation. New Directions for Evaluation* (Vol. 2015, pp. 31–42).

Morris, M., & Clark, B. (2013). You want me to do WHAT? Evaluators and the pressure to misrepresent findings. *American Journal of Evaluation*, *34*(1), 57–70.

Morris, M., & Cohn, R. (1993). Program evaluators and ethical challenges: A national survey. *Evaluation Review*, *17*(6), 621–642.

Morris, M., & Jacobs, L. R. (2000). You got a problem with that? Exploring evaluators' disagreements about ethics. *Evaluation Review*, *24*(4), 384–406.

Neuendorf, K. A. (2016). *The content analysis guidebook: Second edition*. London: Sage publications.

Newman, D. L., & Brown, R. D. (1996). *Applied ethics for program evaluation*: Sage.

Pleger, L., & Sager, F. (2016a). 'Don't tell me cause it hurts'-Beeinflussung von Evaluierenden in der Schweiz. *Zeitschrift für Evaluation*, *15*(1), 23–59.

Pleger, L., & Sager, F. (2016b). Die Beeinflussung in der Evaluationstätigkeit in der Schweiz und was die SEVAL dagegen tun kann. *LeGes-Gesetzgebung & Evaluation*, *27*(1), 33–49.

Pleger, L., & Sager, F. (2016c). Betterment, undermining, support and distortion: A heuristic model for the analysis of pressure on evaluators. *Evaluation and Program Planning*, *early online: http://dx.doi.org/10.1016/j.evalprogplan.2016.09.002*.

Pleger, L., Sager, F., Morris, M., Meyer, W., & Stockmann, R. (2016). Are some countries more prone to pressure evaluators than others? Comparing findings from the United States, United Kingdom, Germany, and Switzerland. *American Journal of Evaluation*. (Early online: 1098214016662907).

Pope, K. S., & Vetter, V. A. (1992). Ethical dilemmas encountered by members of the American Psychological Association: A national survey. *American Psychologist*, *47*(3), 397.

Raupp, J., & Vogelgesang, J. (2009). *Medienresonanzanalyse: Eine Einführung in Theorie und Praxis*. Wiesbaden: Springer VS.

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, *80*(1), 1–22.

SEVAL standards. (2017). Evaluation Standards. Retrieved from http://www.seval.ch/documents/Standards/SEVAL-Standards%202016_d.pdf

Stockmann, R., Meyer, W., & Schenke, H. (2011). Unabhängigkeit von Evaluationen. *Zeitschrift für Evaluation*, *10*(1), 39–67.

Stufflebeam, D. L. (1994). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *Evaluation practice*, *15*(3), 321–338.

The LSE GV314 Group. (2013). Evaluation under contract: Government pressure and the production of policy research. *Public Administration*, *92*(1), 224–239.

Turner, D. (2003). Evaluation ethics and quality: Results of a survey of Australasian

Evaluation Society members. *AES Ethics Committee, November: http://www. aes. asn.

au/images/stories/files/About/Documents*.