

This file was downloaded from BI Brage,  
the institutional repository (open access) at BI Norwegian Business School  
<http://brage.bibsys.no/bi>

## **The dark side of leadership development: An exploration of the possible downsides of leadership development**

Jan Ketil Arnulf, Lars Glasø, Anette K. B. Andreassen and Øyvind Lund  
Martinsen

This is the article as published in

***Scandinavian Psychologist*, (2016) 3: e18**

DOI: <http://dx.doi.org/10.15714/scandpsychol.3.e18>

This is an open access journal available at <http://psykologisk.no/sp/> Articles can be deposited in the institutional repository provided its original source is cited clearly.

# The dark side of leadership development: An exploration of the possible downsides of leadership development

---

 [psykologisk.no/sp/2016/12/e18/](https://psykologisk.no/sp/2016/12/e18/)

**Our interview and questionnaire study shows how leadership development may not be unequivocally positive but implies risks to both participants and organizers if not carried out properly, write Jan Ketil Arnulf and colleagues.**

BY: Jan Ketil Arnulf, Lars Glasø, Anette K. B. Andreassen and Øyvind Lund Martinsen

Research on the effect of leadership development is usually focused on the breadth of positive effects ([Avolio, Avey, & Quisenberry, 2010](#); [Collins & Holton, 2004](#)). To our knowledge, no one has thus far explored whether leadership development may actually have downright negative effects. Negative effects are also unlikely to be detected by prevalent evaluation practices. The purpose of this study is to create a framework for understanding how leadership development could possibly have negative effects on participants and then to empirically explore the existence and prevalence of such effects, taking evaluation practices into account. Even if negative and counter-productive phenomena in organizations have received substantial attention in recent years ([Vardi & Wiener, 1996](#)), we have not found other studies that have explored whether leadership development may actually have downright negative effects.

Leadership development is a multi-billion-dollar industry globally and one of the most expensive activities in corporate training budgets ([Fulmer & Goldsmith, 2001](#); [Yukl, 2012](#)). Specifically, more than USD \$12.5 billion is spent on leadership development, amounting to one-fourth of the total volume of training and development in organizations ([O'Leonard, 2010](#)). The term 'leadership development' refers to a wide range of different activities and the actual impact of these activities on leadership and organizational performance is not widely researched. Furthermore, that impact is even less regularly evaluated by organizations. Collins and Holton published one of the very few meta-analyses in 2004, and they concluded that research suggested "very large positive outcomes, or no outcomes at all." Some more rigorous evaluations of specific interventions ([Antonakis, Fenley, & Liechti, 2011](#); [Collins & Holton, 2004](#); [Feldman & Lankau, 2005](#); [London, 2002](#); [Salas, Rozell, Mullen, & Driskell, 1999](#); [Stoltz, 1992](#)) support this finding. Such contrasting outcomes may be due to different applied practices as well as different methodological designs and measurement methods, different samples, and dissimilar contextual and cultural frames of references. Contrasting results have led some scholars to suggest that leadership development may be grounded in management fads rather than in knowledge about effective approaches ([Collins, 2000](#); [Strang & Macy, 2001](#)).

Existing research indicates that leadership development activities may sometimes be ineffective ([Avolio et al., 2010](#)). The question of whether such activities may even have detrimental consequences for participants comes from a comparison with psychotherapy research, which shows that despite all intentions of doing good, popular interventions like personal feedback and confrontation may sometimes not be in the best interest of the client. One reason average treatment results sometimes appear insignificant may actually be due to negative effects for some participants. The real effects are thus hidden behind differentiated effect patterns ([Barlow, 2010](#)). Closer inspection of unsuccessful treatment has revealed that some interventions have been applied unprofessionally with harmful rather than beneficial or indifferent effects ([Barlow, 2010](#)). Scandinavian media outlets have also reported incidents where managers suffered breakdowns or even attempted suicide after participating in leadership development activities such as sensitivity training ([Hedaa, 1997](#)). (See also the Norwegian newspaper *Verdens Gang* 10.02.1975, 14.06.1997, and 28.10.2000.)

To the best of our knowledge, no existing research has addressed the question of adverse effects of leadership development. Hence, the present study aims to examine the prevalence and types of negative effects that may occur among participants in such activities. We believe this inattention is related to an undeveloped evaluation practice in organizations, and therefore we include this aspect in our design. The contribution of this study is

threefold. First, we present some theoretical arguments on why negative effects of leadership development interventions (LDI) may be likely to occur. Second, we gather empirical data regarding the prevalence and kinds of negative effects that can be related to LDI. And third, we investigate and discuss the consequences of organizations' evaluation practices related to LDI.

### Theoretical background

Leadership development has been defined as: "...the expansion of a person's capacity to be effective in leadership roles and processes" (McCauley & Van Velsor, 2004, p. 2). It has also been described as "every form of growth or stage of development in the life cycles that promotes, encourages, and assists the expansion of knowledge and expertise required to optimize one's leadership potential and performance" (Collins & Holton, 2004, p. 218). As with leadership, no unified definition of leadership development exists. That is because the definition often varies depending on context and focus (Arnulf, 2014). In the present study, we lean on McCauley and Van Velsor (2004) who define "leadership development interventions" as any activity that is intended to expand and improve leadership capacities. The mainstream management literature frequently makes a differentiation between leader development and leadership development. *Leader development* is thought to occur primarily through training that focuses on the individual leader's intrapersonal skills and abilities. A complementary perspective labelled *leadership development* approaches leadership as a social process that engages everyone in the organization. While leader development focuses mainly on individual skills, such as self-awareness, self-regulation, and self-motivation, leadership development focuses more on relational skills, such as social awareness, social skills, and team orientation (Day, 2001).

In the present study, we acknowledge this theoretical difference but use the term "leadership development" to denote both approaches and perspectives. Leadership development activities may be divided into three types of activities: 1) general management education, such as undertaking an (E)MBA program or learning formal skills; 2) internal organizational activities aimed at creating strategic alignment and social capital, such as kick-offs and strategic planning seminars; and 3) individually targeted development activities directed at building and improving generic leadership capabilities ("leader development") (Arnulf, 2014).

Based on prevalent definitions of leadership development, we therefore define "negative effects" of leadership development as follows: Leadership development activities have negative effects if they, contrary to intentions, either (a) directly reduce the person's capacity to perform leadership roles or (b) indirectly reduce organizational performance by wasting resources and undermining the belief in developmental efforts.

A common complaint against individual "leader" development is that traditional teaching-style interventions such as MBA programs have had a tendency to deliver declarative or theoretical knowledge but seem less effective in creating "actionable" knowledge (Argyris, 2000; Baruch, Bell, & Gray, 2005; Baruch & Leeming, 2001; Benjamin & O'Reilly, 2011; Khurana, 2007; Livingston, 1971; Mintzberg, 2004). A prevailing opinion shared by practitioners and researchers holds that the success of individual leaders depends on personality characteristics (Quirk & Fandt, 2000) and other non-academic characteristics that are robust enough to make people speculate whether leadership is actually innate rather than learned (Avolio, 2004). Given this focus on individual non-academic skills, many approaches to leadership development target self-awareness (e.g., by providing feedback that challenges self-serving defense mechanisms) (London, 2002; McCauley & E. Van Velsor, 2004). Interventions that depart from formal education on leadership targeting tacit and person- (if not personality) based abilities frequently prompt leadership developers to turn to clinical psychology to find rationales for their interventions (e.g., Berglas, 2002; Kets de Vries & Balazs, 2005). Practitioners are often eager to point out the differences between leadership development and clinical psychology, but such statements may be more rhetorical than real for three reasons. First, clinical phenomena such as defensiveness, twisted reality perceptions, and hidden personal agendas will always affect the psychology of leadership (Dixon, 1976). Second, the assumption of a clear division between pathology and normality defies the epidemiological fact that clinical conditions are too common to ignore in working life (de Graaf, Tuithof, van Dorsselaer, & ten Have, 2012). Third, the broad aim of leadership development in improving "capacities for filling leadership roles" usually goes beyond the mere teaching of simple skills (Day, 2001; London, 2002; Schein, 2006; Takala, Winegar, & Kuusela, 2009). For this reason, numerous leadership development activities either themselves or as programs resemble psychotherapeutic interventions (Berglas, 2002). This resemblance includes helping the participants learn and understand what they feel and what others feel about them. It also encompasses learning to be more sensitive to the effects of their behavior on others and, accordingly, learning to develop specific behavioral interpersonal

skills. A systematic study of leadership development's effectiveness may therefore parallel the study of psychotherapies in interesting ways. Evaluation research in clinical psychology has documented differential treatment effects, often running counter to the practitioners' original assumptions and sometimes including unintended negative outcomes (Barlow, 2010). The more leadership development interventions move beyond the teaching of management skills and target personality dispositions instead, the more they may suffer from the same quality problems that affect other fields of behavioral change.

Our claim is that evaluation research from clinical psychology offers plausible reasons to theorize about possible negative effects in leadership development. We suspect that such negative effects can be linked to three types of situations: a) ineffective interventions that leave participating managers frustrated because of the waste of time and resources; b) adverse effects on the participants with short- or long-term negative effects on their capability for work and private well-being; and c) a lack of evaluation that contributes to unprofessional practices and obscures the prevalence of negative effects. In the following pages, we will discuss the theoretical assumptions in more detail. We will then move to an empirical exploration of the proposed phenomena.

### **Lack of effect**

Psychotherapy research is a showcase of how well-intended and even widespread interventions face great difficulty in establishing empirical support, a situation that is paralleled by research on leadership development (Avolio et al., 2010; Barlow, 2010; Collins & Holton, 2004; Roth & Fonagy, 1996). There may be several reasons why leadership development fails, such as poor learning and training conditions, unclear learning objectives, meaningless content, misappropriate sequencing of content and mix of training methods, lack of opportunities for active practice, and lack of relevant, timely feedback and appropriate follow-up activities. According to Gurdjian, Halbeisen, and Lane (2014), leadership development fails because contextual factors are not taken into consideration, reflections are decoupled from the real world, the effects of participants' mindsets are underestimated, and results are not measured (Gurdjian, Halbeisen, & Lane, 2014). A meta-analytic review of experimental and quasi-experimental studies (Avolio, Reichard, Hannah, Walumbwa, & Chan, 2009) found that leadership development had moderate positive effects on training groups compared to controls. However, in their meta-analysis Powell and Yalcin (2010) investigated the effects of managerial training from 1952 to 2002. They found significant (positive) effects of such training during the 1960s and 1970s but not during the 1950s, 1980s, and 1990s. As such, these results are mixed and complicated to explain. Another meta-analysis, this one focusing on the transfer effects of training, suggests that interpersonal skills are particularly resistant to change, with very low transfer effects from the training to perceived differences by peers and subordinates (Taylor, Russ-Eft, & Taylor, 2009). For example, a study of overall effect size estimates of the transfer of managerial training showed a pattern reflecting the importance of the rating source. Such effects changed significantly depending on the rating source (e.g., rating from self-report: 0.64; superior: 0.53; peer: 0.26; and subordinate: 0.13) (Harms, Spain, & Hannah, 2011). Research on the effects of leadership development has shown mixed findings, reporting both positive and negative effects. A non-trivial point here is that there may be within-individual variation as well. The same individual may have experienced some interventions as helpful and others as less so. This tendency can cause irritation because the activity does not address the needs of the individual and the organization. Thus, it may be perceived as a distraction or waste of resources. Information about both positive and negative experiences from the same subjects is important both to document the credibility of respondents and to give an estimate of how managers may be negatively and/or positively biased in their responses. We therefore formulate our first hypothesis, which will be explored empirically:

- *Hypothesis 1: Managers will report both positive effects and a lack of effect from participation in leadership development programs.*

### **Short- and long-term adverse effects**

Most professionals working in close relationships with clients (lawyers, medical doctors, psychologists, clergy, nurses and social workers) to facilitate important life challenges have recognized the need for ethical guidelines since the occupational role of these professions also opens up the possibility for abuse (e.g., American Psychological Association, 2010). The ethical guidelines cover a typical range of malpractice risks such as a) inappropriate choice of interventions that exacerbate rather than ameliorate the clients' problems; b) accepting assignments without proper qualification and abusing the client's situation to satisfy the helper's own needs (thereby creating dependency, sexual relationships, or other transgressions of the client's integrity); or c) using information that was submitted in what seemed like an intimate and trusting relationship for purposes that violate



the interest of the client or third party. The case against Rajat Gupta (the former CEO of McKinsey) for leaking inside information from clients is a reminder that this scenario may also unfold in other types of helper/client-relations (Hill, 2011).

Taking part in leadership development programs is an offer that is difficult to turn down for managers. That is because the invitation to such programs signals career opportunities for the invited. When participants choose to disclose personal information and commit themselves to the processes of such programs, they may assume that this is for their own good. However, employers, other participants, and external consultants alike may use personal information about the participants in ways that run counter to their interests and even violate their personal integrity. This may involve blocking promotions, disclosing personal matters, or using personal information for other purposes in ongoing power struggles within the organizations.

Inviting or nominating someone for a leadership development intervention is therefore likely to bring together a series of conditions that: a) trigger emotional engagement because the intervention is seen as targeting important opportunities and obstacles in the participant's career progress; b) increase the risk of failure to perform because of the organization's investment in management performance evaluation; and c) induce the facilitator – whether internal or external – to apply dramatic procedures in order to prove the intervention's potential value to the involved stakeholders. When effective, such conditions may create a sense of urgency, personal involvement, and organizational relevance, as is typically wanted in procedures such as 360-degree feedback (DeNisi & Kluger, 2000; Kluger & DeNisi, 1996; Taylor & Bright, 2011). The most common risk is probably that none of these byproducts occurs and the event passes as unimportant, but the situation may backfire on participants who become involved but also exposed and possibly feel betrayed or outcast if the process fails to meet their expectations. While the ability to learn from stressful events is partly dependent on personality factors (Cooper, 2010), we have rarely heard of leadership development programs that match learning approaches to individual stress-coping diagnostics, and such mismatch is therefore likely to appear in larger samples of participants.

In this paper, adverse effects include painful and/or harmful experiences for the target person to an extent that was not justified by perceived gains in performance. This leads to our second hypothesis, which will be examined in Study 2:

- *Hypothesis 2: Instances of adverse experiences will be reported in leadership development, with some cases involving long-term emotional distress.*

### **Lack of evaluation**

Proper evaluation of leadership development is possible but methodologically challenging. This type of knowledge is not likely to be present in most organizations that engage in or buy leadership development (Pfeffer & Sutton, 2006). Even when LDIs are evaluated, the most common methods of evaluation are likely to be measures of participant satisfaction, not learning outcomes or organizational benefit (Aragón-Sánchez, Barba-Aragon, & Sanz-Valle, 2003; Kirkpatrick, 1979; Van Buren & Erskin, 2002). Moreover, there is no guarantee of improved organizational performance. Even if there are some guidelines that may improve the outcome of LDIs (Mabey & Finch-Lees, 2008; Whetten & Cameron, 2005), it remains an empirical fact that organizations probably conduct leadership development without systematic evaluation, as shown in a literature search for meta-analyses in this field (Collins & Holton, 2004). We also suspect that facilitators from within the organization are more sensitive to the local needs and more accountable to their organizations than externally hired consultants are.

Moreover, we believe that an interest in evaluation will signal that the organization is following a reflected and informed approach to leadership development as a way to reach defined targets. Such an instrumental approach to development is more likely to follow a rationale of personal and organizational diagnostics and meaningful pedagogical design, as recommended by research (Collins & Holton, 2004). Conversely, we assume that lack of evaluation co-exists with less focus on the design of activities, hence increasing the likelihood of unwanted outcomes. We therefore formulate the following two hypotheses:

- *Hypothesis 3a: Leadership development programs are frequently conducted without systematic evaluation.*

- *Hypothesis 3b: Instances of negative or ineffective leadership development will be found to co-exist with a tendency to abstain from evaluation of leadership development activities.*

We know from initial pilots of this study that there are widely different practices and understandings of how leadership development may be evaluated in various organizations. Within this study, we are more concerned with the intentions to evaluate than to explore the specific methods being used. The true prevalence of negative LDI outcomes is probably difficult to assess. For one, employers most often initiate and pay for LDIs. Managers take part because they want to improve their career chances and thus will be reluctant to criticize or expose their vulnerability. There are no quality standards or guidelines against which the performance of facilitators in leadership development may be matched (Feldman & Lankau, 2005). Complaints are therefore likely to be futile. Furthermore, those responsible for the activity may be reluctant to criticize openly, lest they undermine an already disappointing investment. Lastly, a lack of evaluation competence among practitioners may block exchange of experience and render the whole issue obscure. For such reasons, there is no easy and direct way to estimate the prevalence of possible negative effects. To handle this problem, we explore the field by employing a triangular empirical approach in three studies:

Study 1 is a qualitative set of interviews with people who share first-hand experiences with cases of negative effects. This study was conducted to explore the question of *how* such negative effects appear and are perceived by the affected, with no claims about how frequently such incidents occur.

Study 2 is a survey of a wide range of managers from different companies and businesses in Norway that are trying to establish the *prevalence* of such experiences, as examined in the interviews in Study 1.

Study 3 is a similar survey on the evaluation practices in companies that conduct leadership development. In order to obtain a balanced set of data, we have some overlap among the companies in studies 2 and 3, such that some managers have responded only in study 2, some have responded only in study 3, and some have responded in all three studies. In this way, we establish an estimate of evaluation practices. We also link these practices to the negative outcomes that are central to our study.

## Study 1: Exploring types of negative experiences

### Methods

First, we sought to qualitatively explore clear-cut cases of malpractice where the participants were obviously experiencing harmful effects from the LDI experiences. We therefore sought to find people who could substantiate claims about problems related to somatic and mental health as well as social relationships and the context where such claims arise.

**Procedure** – We browsed a broad range of library databases such as PsychLit, Web of Science, and Internet search engines to look for documentation of harmful effects of leadership development. We also contacted professional organizations such as the Norwegian and the American Psychological Associations to find out if they had been discussing or addressing complaints about such topics, but we found no systematic sources documenting negative or harmful effects of leadership development. As far as official sources are concerned, it seems as if leadership development never implies any risks other than having no effect. The search for legal cases, both nationally and abroad, also yielded no results. We contacted legal professionals in case they knew about potential cases that had been framed as bullying or had been handled internally in the organizations as breaches of their code of conduct. We also spread the message via our social collegial networks. Finally, we posted an advertisement in leading Norwegian newspapers asking for information from people who had first-hand experience with negative effects of LDI.

**Sample** – This approach yielded a sample of 14 participants. Of these, seven were clients who had been negatively affected by leadership development. Another seven were professionals who had treated people with negative reactions after such activities (clinical psychologists), law professionals, or union leaders who have been giving advice to clients and who felt they were subject to unethical practices at work. Two persons actually appeared in both roles. They were themselves practitioners of leadership development who had entered into

activities they felt a need to warn others about. All were invited to a semi-structured interview (Kvale & Brinkmann, 2009). The structure consisted of us presenting our general ideas about the topic—also the reason for conducting the interview—whereupon we asked the subjects to describe their experiences as richly as possible. The subjects not only answered our questions and elaborated on them, but they also gave us access to notes taken and actual formal reports handed in to authorities to warn about certain practices. This spirit of openness resulted in very rich and complex material spanning phenomenological, clinical, and juridical viewpoints on our topic (Kvale & Brinkmann, 2009).

Analysis of the interviews was performed by two of us in a bottom-up way, according to the principles of grounded theory (Glaser & Strauss, 1967). The preliminary categorization resulted in three overarching categories of events: A) psychological distress or psychopathology connected with the activity; B) abuse of advisor-client relationship; and C) unethical use of information. These categories were coded with four levels of impact: 1) negative but relatively harmless (spoiling positive benefits that may have been reached by other approaches); 2) creating a transient painful experience with no beneficial effect; 3) leaving painful after-effects; and 4) having detrimental consequences such as suicide, permanent health damage, or loss of job.

### Results from the interviews

In the following passages, we quote and list findings from the sources described above. These are all statements from individuals except from some media excerpts. Such instances are noted as sources.

**A) Addressing or eliciting distress or psychopathology** – Several statements illustrate this category. For example, *“He (the consultant) was rather surprised when the conflicts came up... I believe he was quite poorly prepared actually”* is illustrative of an ineffective but, as we see it, rather harmless negative effect of consultant unawareness. On the other hand, the statement *“Firstly, we had a session that concerned getting to know one another. One quit, because ... [s/he] received a lot of unwelcome comments”* refers to a more painful experience. In contrast, an example of a painful after-effect would be: *“I became a symbol for what they wanted and did not want. It was a very difficult workplace. I started crying: this I cannot take.”*

The category *Use of group pressure* comprises painful experiences such as *“Colleagues broke down and were provoked into frustration. Woman completely shutting down and withdrawing herself from the seminar.”* Examples of experiences presuming painful after-effects were *“Stress-reactions, anxiety, trouble sleeping, numb feeling in body”* and *“Still have physical reaction from seeing the consultant.”* Comparatively speaking, the impact of statements like *“Colleague with known psychological illness had breakdown and [was] sent home in taxi; later he committed suicide,”* and *“Casualties; temporary psychosis and short episodes of disorientation during seminar; problems with identity, manic-depressive episodes”* were understood as having detrimental consequences to the participants.

Regarding *Targeting personality-based behavior*, the participants' experiences seemed to have a different impact on the targets. Ineffective but relatively harmless experiences were reported, such as *“Confronting prejudice and social phobias by performing acts they were fearful/negative of (e.g., cross-dressing and going to a gay bar to cure homophobia).”*

Examples of painful experiences included *“Improper exposure of personality test results: Had I not been working in a snug and close work environment, this could have gone truly wrong.”* The experience and statement *“Confusion, crying, stress, lack of sleep: To me, this elicited a large quantity of processes, and I was fairly sidelined for a long while afterwards”* is seen as causing painful after-effects, while experiences related to statements like the following (all three of which are excerpts from newspaper articles) are understood as having detrimental consequences for the participants: *“Leader suffered breakdown after [being] put in isolated position based on personality testing,”* *“Director of transport company in Norway shot himself after a result of an organizational analysis,”* and *“Top-leader suicide in Denmark after an evaluation.”*

**B) Abuse of advisor-client relationship** – Statements of unethical relationships with abuse of power were reported too. Examples of incidents with relatively harmless effects included the following: *“Consultants talking negatively about the superiors of participants, glorifying own methods while talking negatively of others' ways.”* Meanwhile, statements such as *“He took a thorough time to think, then he looked around and pointed at me saying, ‘You are the only person who has asked this question!’”* are seen as having a painful impact on the

participant. One informant spoke of painful after-effects this way: *"I went home believing I was a manipulative person."* The reviewed media sources (Hedaa, 1997, see also *Verdens Gang* 02.1975, 14.06.1997, and 28.10.2000) also reported detrimental consequences, such as *"Participant feeling like a failure in work and in his family. Was not sure life was worth living."*

**C) Unethical use of information** – The category *Unethical use of information* contains three sub-categories. One is the breach of participants' confidentiality, including statements like *"It became a damned racket, because they had said that it would be anonymous when they interviewed the participants. And then we got the papers handed out... quoting specific people on specific issues and so on."* Another example is the following: *"This person has still not forgotten about it. So in tears she cried: 'If I had known, I would not have told anyone.'"* Another sub-category is labeled *Information leaked to superiors*. Here, an informant reported that *"[Not even a week went by] until I was confronted with things I knew had to have come from me or him. And I knew that I had not said them in public"* and *"Consultant uses personal information from group processes against participants later on."* Examples of statements with painful after-effects were *"Coaches being 'Information octopuses.' Instances of bullying. Mental strain and anxiety."* Finally, some statements seemed to fit under a sub-category labeled *perceived intervention*. An example of a statement with detrimental effects was the following: *"Participants informed by superiors that they would not have a job if they quit the development program. Massive sick-leaves and calls for psychiatric help."* The analysis of the interviews is summarized in Table 1.

**TABLE 1:** Summary of negative effects of leadership development activities as reported by informants. Sources coded as: C = Case informant, P = Professional informant, M = Media / reports / literature.



Type of negative experience:	Degree of impact: Negative but relatively harmless	Degree of impact: Painful experience	Degree of impact: Painful after-effects	Degree of impact: Detrimental consequences
<b>Addressing or eliciting distress or psychopathology</b>				
Consultant unawareness	"He (the consultant) was rather surprised when this (the conflict) came up .... I believe it was quite poorly prepared actually." (C)	"Firstly, we had a session that concerned getting to know one another. One quit, because... [s/he] received a lot of unwelcome comments." (C)	"I became a symbol for what they wanted and did not want. It was a very difficult workplace. I started crying, 'this I cannot take.'" (C)	
Use of group pressure		Anxiety symptoms like groupthink, worship of leaders, and power struggle with aggression or submission (C)  Colleagues breaking down and provoked into frustration; Woman completely shutting down and withdrawing herself from the seminar (C)  CEO sacks department managers in frustration. Did not act upon this later on. (C)	Stress reactions, anxiety, trouble sleeping, numb feeling in the body (C)  Exhaustion, fatigue, sleeplessness, paranoia (C)  "Personal stories were shared in ways that most of us still struggle with afterwards." Weight loss, sleeplessness, stopped eating, insecurity, depression (C)  Still have physical reaction from seeing consultant (C)  "One had to be held back [so as] not to jump out of the window." (C)  Casualties; Temporary psychosis and short episodes of disorientation during seminar; Problems with identity, manic depressive episodes (P)  ZZZ-Denmark and [supplier company]: Psychosis and hospitalization after development seminar (M)	Two suicides related to confrontational group processes at XXX-seminars (P)  Suicide in 1974 in Sweden associated with sensitivity seminar related to work (M)  Facilitator from third-party consulting company causing mental trauma, depression, job quitting, sick leave, etc. (M)  Sick leave, psychiatric aid. "I was a visible skeleton to many. I was very upset." Several coworkers also received the same assistance. (C)  Colleague with known psychological illness who had breakdown and was sent home in a taxi, later committed suicide. (C)
Targeting personality-based behavior	Confronting prejudice and social phobias by performing acts they were fearful of / negative about (e.g. cross-sex dressing and going to a gay bar to cure homophobia) (C)	Improper exposure of personality test results: "Had I not been working in a snug and close work environment, this could have gone truly wrong." (C)	Confusion, crying, stress, lack of sleep: "To me, this elicited a large amount of processes, and I was fairly sidelined for a long while afterward." (C)  Leader suffered breakdown after being put in isolation based on personality testing. (P)	Director of transport company in Norway shot himself following organizational analysis (P)  Top-leader suicide in Denmark after an evaluation (P)
<b>Unethical use of relationships</b>				
Abuse of power	"We once had to break off the session when he and I clashed for the first time, get the others out ... because he (the consultant) massively attacked me as a person, you know." (C)  "When I reply and he cannot take it, then I am sort of diagnosed as a confirmation of all the bad things he is saying." (C)  Consultants talking negatively about the superiors of participants, glorifying their own methods while talking negatively of others' ways (C)	Consultants pushing and bullying vulnerable participant having gone through surgery (C)  Consultant slapping shut lap-top over participant's fingers in order to make an example for those not paying attention (C)  Branded a troublemaker after questioning the consultants: "He took a thorough time to think, then he looked around, and then he pointed at me saying, 'You are the only person who has asked this question.'" (C)  "I guess others felt more uncomfortable than we did. Some felt harassed, and it is pure harassment in the presence of others one is subjected to." (C)  Consultant uses	Systematic psychological breakdown: "I went home believing I was a manipulative person." (C)	Participant feeling like a failure in work and in his family; was not sure life was worth living (M)

		personal information from group processes against participants later on. (C)		
<b>Unethical use of information</b>				
Breach of participant's confidentiality	<i>"It became a damned racket, because they had said that it would be anonymous when they interviewed the participants. And then we got the papers handed out ... quoting specific people on specific issues and so on." (C)</i>	<i>"This one person has still not forgotten about it. So she broke down in tears: 'If I had known, I would not have told anyone.'" (C)</i>		
Information leaked to superiors	<i>"And not even a week went by until I was confronted with things I knew had to have come from me or him. And I knew that I had not said them in public." (C)</i>		Coaches being "Information octopuses"; Instances of bullying; Mental strain and anxiety (M) (P)	
Perceived intervention		<i>"One would mostly get fixated on one or several people whom one wanted to challenge. And was allowed to attack." (C)</i>	<i>"Her superiors consented when it actually happened. No one protested. I believe it has affected her quite strongly, actually." (C)</i>	Participants informed by superiors that they would not have a job if they quit the development program: Massive sick leaves and calls for psychiatric help (C)  Head of division fired after receiving coaching (P)

According to the interviewees, many of the theoretical possibilities for negative incidents had actually occurred. They expressed concern for the competence of the consultants they had come across, often experiencing that processes had gotten out of control. The external consultants were often described as overly enthusiastic about their own methods, being blind to alternative perspectives/approaches, and being unaware of contextual limitations such as organizational and individual needs. In addition, the use of group pressure and therapeutic elements led participants to feel as though the consultants did not seem to be aware of the implications and contingencies of such experiences. Our subjects described some consultants as both psychologically and physically intrusive and as having shared confidential information with others in the organization (most notably employers) while manipulating those who raised doubts about the value of the ongoing activities.

Thus, we think our interviews show that professionals engaging in leadership development are susceptible to the same ethical and technical risks other professionals who are engaged in close relationships with clients face. Hypotheses 2 and 3a are thereby supported. We would like to add, however, two notable possibilities that were mentioned indirectly (had been heard of /observed) but were not reported as personal experience in our interviews: 1) consultants engaging in sexual relationships with clients and 2) instances of inside information being used for financial transactions.

## Study 2: The prevalence of negative effects

### Methods

**Procedure** – A total of 214 questionnaires were distributed to a diverse convenience sample consisting of 99 managers participating in executive education and 115 participants in an HR conference. Of these, 189 respondents completed and returned responses, yielding a response rate of 88%.

**Sample** – The sample consisted of 52% men and 48% females. The age range was 18 to 66 years. Age was listed in intervals of 10 years, a tendency that did not allow for exact statistics, but both mode and mean were in the age group of 36 to 45. The distribution of managerial experience was senior management position 18%, middle management 54%, other management position 16%, and no managerial experience 12%.

**Instruments** – The questionnaire was designed to ask about the prevalence of the experiences explored in Study 1. We think there is little sense in quantifying each category in terms of intensity or times of experience, primarily because we lack reliable information about this frequency in the respondents' memories and secondarily because doing so may create an unwieldy number of categories with scant statistical value. Instead, we simply listed the following types of leadership development activities: 'Developmental seminar for personnel';

'Teambuilding activities for personnel'; 'Individual coaching'; 'Development program / seminar for managers'; and 'Teambuilding activities for management teams.' For each activity, we asked the respondents if they had ever experienced these as 'Positive experiences,' 'Wasted experiences,' 'Negative experiences,' or 'No such experiences.' Multiple responses were allowed. Our numbers were obtained by counting the number of respondents who rated their experiences for each alternative. Asking people for both positive and negative experiences of the same types of activity was important to check the sample's overall negative or positive attitude.

## Results

As shown in Table 2, almost all participants had experienced leadership development activities as positive. Interestingly, two-thirds of the same group had experienced such activities as wasted, and almost half the sample reported negative experiences with such activities. (Note that since it is possible for respondents to have experienced more than one activity in both positive and negative ways, the percentages may sum up to more than 100%.) Hypothesis 1 is thereby supported.

**TABLE 2:** Frequencies of positive, wasted, and negative experiences of developmental activities.

Activity	Positive experience %	Experienced as wasted %	Negative experience %
Developmental seminar for personnel	60.8%	36.5%	16.0%
Teambuilding activities for personnel	70.4%	33.9%	12.8%
Individual coaching	48.1%	12.2%	8.0%
Single-session leadership development	46.6%	24.3%	0.0%
Development program seminar for managers	44.4%	11.6%	10.6%
Teambuilding for managers	44.4%	21.2%	10.6%
No such experience	2.6%	37.0%	62.2%

The results showed that activities including a larger number of people (e.g., seminars and team building) were more likely to be experienced as wasted and/or negative than individually targeted activities such as coaching. Table 3 shows the reasons for wasted experiences and indicates that the respondents seemed to blame the wasted time on low-quality facilitators.

**TABLE 3:** Reasons given for why activities were experienced as wasted ( $N = 120$ ).

The perceived reasons for negative experiences are presented in Table 4. Here, the common denominator seems to be that participants are exposed to situations they feel are embarrassing or intrusive and those for which they did not feel prepared.

**TABLE 4:** Reasons for negative experiences and the severity as a percentage of the 69 respondents who listed such experiences (37% of the total sample, multiple responses allowed).

Reason why this was wasted time	Percentage ( $N = 120$ )
Activity not adapted to my needs	45%
Learned nothing useful in my job	55%
The facilitator was not doing a good job	55%
There was no follow-up	31%

Note: 69 (37%) did not have such experiences, and multiple options were allowed.



Reason for negative experience	Percentage (N = 69)
The activity turned into something I was not prepared for.	57%
I did not feel that the facilitators respected my integrity.	36%
I felt forced to say or do things that made me feel awkward.	36%
Information that I shared during the activity was used against me later.	28%
I experienced feedback from the facilitator(s) as insulting.	20%
I experienced feedback from other participants as insulting.	14%
This activity harmed my ability to carry out my work.	6%
Severity of effect afterward	
I quickly got over it.	57%
I was emotionally upset for a while.	36%
It has negatively affected my workplace relations.	17%
It negatively affected my private life.	9%

Of those who reported negative experiences, a small majority put this negativity quickly behind them. Still, 25 respondents felt the effects for a while, affecting both their private life and their relationships in the workplace. Of the 69 respondents who had had negative experiences, 12 (17%) held a co-worker responsible, 21 (30%) thought their superior manager was responsible, and 53 (77%) thought an external consultant was responsible for the negative experience. Of the sample, 62% claimed to have heard about others who were negatively affected by LDI activities. Hypothesis 3a is thereby supported.

### Study 3: Prevalence of evaluation practices

#### Methods

A questionnaire was sent to 116 training professionals who were working with leadership development in organizations, yielding 50 complete responses. Additionally, 93 of the HR professionals taking part in Study 2 were also asked about their company's engagement in leadership development activities and their evaluation practices.

**Measures** – To capture the level and effort invested in evaluation practices, we created a survey based on Kirkpatrick's four-level model (1979) since this is the framework most frequently used by practitioners (Aragón-Sánchez et al., 2003; Arthur, Bennett Jr., Edens Jr., & Bell, 2003) and also includes both individual developmental outcomes as well as organizational-level outcomes. The framework increases in complexity according to levels, requiring more competence for higher levels of evaluation (Kirkpatrick, 1979):

1. Level 1 concerns the participants' "affective and attitudinal response to the training program." This response is often used by organizations that assume positive reactions increase the probability of successful learning outcomes, which has not been shown to hold empirically (Arthur et al., 2003).
2. Level 2 concerns the measurement criteria for learning outcomes. Kirkpatrick defines learning as the "principles, facts, and techniques [that are] understood and absorbed by the conferees" (1979, p.80). He suggests two methods for assessing learning outcomes: a practical evaluation and a written test.
3. Level 3 concerns the transfer of training. Here, the learning attained and measured at level two is tested for application (i.e., the possibility to develop "know-how" and thus competitive advantage for the organization).
4. Level 4 concerns organizational outcomes and seeks to establish measures by placing a monetary value on the improved quality or increased efficiency that results from learning on Level 3.

Four items addressing evaluation measured the prevalence of evaluation practices on Levels 1 and 2. These benchmarks targeted the participants' immediate reactions through feedback sheets and satisfaction surveys) and tested participants' knowledge after attending the LDI (classroom-testing and final exams). Prevalence of



Level 3 evaluation was measured by items asking for practices such as 360-degree feedback, goal setting, work environment surveys, and on-the-job observations (Hannum, Hannum, & Reinelt, 2007). An example item is: "To evaluate leadership development programs, we are: – using 360-degree feedback... – using on-the-job observations...". Prevalence of Level 4 evaluation was measured by asking if and how frequently the responding companies were calculating return on investment (ROI) for leadership development or were using other objective performance indicators. Additionally, we asked the respondents to rate their perceived competence of the evaluation, their degree of decision-making authority in initiating programs and evaluations, and the approximate sums of money spent on leadership development and evaluation. All items were measured on seven-point Likert scales and were constructed solely for this project. Alpha reliabilities for all scales ranged between .60 and .85, which by convention means a range from questionable to good. However, recent advances in statistics suggest that lower alphas may indicate that the information has been sampled from a broader range of indicators, thus giving a more correct measurement of the phenomenon (Arnulf, Larsen, Martinsen, & Bong, 2014; Kline, 2000). This is particularly so when the objective of measurement is assumed to be formative rather than reflective, as in this case (MacKenzie, Podsakoff, & Podsakoff, 2011).

## Results

The results from the survey evaluation practices are summarized in Table 5. It shows that organizations are evaluating leadership development "fairly often" on Kirkpatrick's Levels 1 and 3 but almost never on Levels 2 and 4.

**TABLE 5:** Mean self-rating of leadership development evaluation practices in organizations.

Evaluation practices according to Kirkpatrick's levels:	Mean frequency rating (1–7)	Std. dev.	N
Level 1 evaluation	4.1	2.0	69
Level 2 evaluation	2.2	1.2	73
Level 3 evaluation	3.6	1.2	67
Return on investment	2.0	1.5	61
Performance indicators	3.0	1.9	67
Total Level 4	2.4	1.5	

Note: Frequency of evaluation practices is rated on a Likert scale from 1 to 7 where 1 is "never" and 7 is "always."

The practices are skewed negatively such that the majority of respondents rated their frequency of evaluation from 1 to 3. Of the entire sample, 33.9% stated that they never evaluate LDIs at Level 4. Only small groups of companies (between 3% and 7%) answered that they "always" evaluate LDIs on every level.

**TABLE 6:** Investments in LDI and in LDI evaluation (percent of respondents).

	% of respondents					
	\$ 0	< USD 100,000	USD 100,000–200,000	> USD 200,000	Don't know	Not relevant
LDI budget	3.3%	16.5%	23.1%	29.7%	27.5%	0%
LDI evaluation						
Budget	39.8%	7.2%	10.8%	7.2%	28.9%	6%

Note: All monetary values are in USD 2010 numbers.

As seen in Table 6, almost all surveyed organizations paid for LDI, but only one-fourth of them could say how much they actually paid for evaluation. Hypothesis 3b is therefore supported. Still, because these organizations actually seemed to have defined expenses expressly for this purpose, the competency in evaluation seems to exist in organizations even if sparsely applied.

The HR professionals taking part in the survey on the negative effects of leadership development were also asked similar questions in order to explore the relationship among professional leadership development, evaluation practices, and the likelihood of creating negative effects. In this case, we had to restrict the number of items to reduce the length of the questionnaire and hence make it less obtrusive. This version asked respondents one-item questions about Kirkpatrick's evaluation levels as well as about the financial and professional resources invested in leadership development programs and their evaluation. The practices are displayed in Table 7.

**TABLE 7:** Distribution of leadership development and evaluation practices among 93 companies, as rated by their HR managers.

Resources invested in company's leadership development					
No leadership development	Only internal resources, no budget	Spend less than NOK 500,000 per year	Spend NOK 0.5–1 million per year	Spend more than NOK 1 million per year	Unknown
1%	9%	35%	17%	22%	21%

  

Types of resources procured in leadership development						
Use internal facilitators with no special qualifications	Use internal facilitators with special qualifications	Use external lecturers	Use external facilitators	Send employees to external leadership development programs	Send employees to leadership education like EMBA	Offer individual coaching
23%	68%	71%	42%	53%	38%	46%

  

How the companies systematically evaluate leadership development					
Never evaluate	Evaluate how participants like each activity	Evaluate participants' subjective knowledge increase	Evaluate participants' objective knowledge increase	Employees rate managers before and after development activities	Evaluate cost / benefit of activities in terms of organizational performance
19%	67%	46%	22%	15%	8%

  

Money spent on evaluation					
Do not spend money on evaluation	Spends less than NOK 100,000 per year on evaluation	Spends NOK 100,000–500,000 per year on evaluation	Spends NOK 0.5–1 million per year on evaluation	Spends more than NOK 1 million per year on evaluation	Unknown
38%	26%	3%	1%	0%	32%

To test hypothesis 3b, we entered the numbers describing company investments in leadership development activities and evaluations from Table 6 into a binary logistic regression. Our goal was to predict the HR managers' positive, wasted, and negative experiences, respectively. These latter values were coded as binary yes / no values for each HR manager if s/he had reported any such experience. These models were significant and strong for positive experiences ( $p < .001$ , Nagelkerke  $R^2 = .41$ ) and significant but weak for negative experiences ( $p = .004$ , Nagelkerke  $R^2 = .20$ ) as well as wasted experiences ( $p = .019$ , Nagelkerke  $R^2 = .16$ ). Hypothesis 3b was thus supported.

## Discussion

The aim of this study was to explore the possible realities behind theoretical considerations and rumors that leadership development interventions may have negative effects. Assuming that gaining empirical access to this phenomenon is difficult, we tried to explore it from three different angles: 1) How do people subjected to negative practices recall their experiences in interviews? 2) What is the prevalence of managers who claim to have experienced positive, wasted, or negative LDIs? 3) What seems to be the professional level of corporate evaluation practices for LDIs, and how does such a level relate to our respondents' negative experiences?

All our hypotheses were supported by the empirical data. While the overwhelming majority of the surveyed managers had experienced developmental activities as positive at least once—and hence cannot be said to have been negatively biased towards this field—a majority of the same sample had experienced LDI as negative but harmless (i.e., a waste of time (78%)) and almost half of them as negative even (44%). The negative experiences that we inferred theoretically were found to occur in practice. The most frequently attributed cause of negative practices was external consultants, operating in an environment characterized by lack of evaluation. The most likely scenario for negative effects to occur seems to be companies that invest quite a lot in the development activities themselves but not in their evaluation. The safest practices seem to be applied by companies that either invest very little and rely on internal resources or invest high sums of money and therefore also monitor the effects closely.

Even if the really tragic outcomes seem infrequent compared to the number of harmless but wasted occasions, we think our study points to what Berglas (2002) has called “very real dangers” of unprofessional or rogue practices in leadership development. In our sample, more than 10% of respondents had experienced negative reactions affecting their private lives and/or workplace relationships for a period extending beyond the event itself. Since almost all the respondents also answered that they had experienced positive instances of LDIs and that they actually believed in the possibility of leadership development, we assume that the frequency of negative experiences was not due to a negative response set.

The subjects who agreed to be interviewed about their experiences testified to the existence of most types of transgressions that also befall other professions purporting to help clients. It certainly seems possible that incompetent tampering with psychopathology, exploiting relationships, and leaking sensitive information to a third party are possible scenarios when engaging in leadership development. Examining the experiences of our interviewed subjects, it becomes obvious that LDI activities can evoke strong reactions in the participants and that these dynamics are very dependent on the qualifications of the facilitators in charge. In particular, external consultants are perceived by many as transgressing professional boundaries with their interventions. Such consultants are the most frequently used facilitators and we have no data to control whether they are more frequently appearing in this role than mere numbers should suggest.

One may wonder why consultants and facilitators engage in these kinds of practices. Are the negative effects unintended side effects and work accidents, or are they sometimes produced on purpose? Our data give no direct answers to such questions, but we may infer some answers from combining the answers. Because evaluation practices rarely reveal LDIs as effective or ineffective, the most salient sign that the consultant is wielding a mighty tool will be the reactions of the participants. Our data suggest that the most likely risk facing an external consultant is that the activity is experienced by participants as ineffective (i.e., a waste of time). From a marketing perspective, this is a risk to the consultant's business, reducing the likelihood of being used again. It is possible that by provoking emotional reactions in the participants, the external consultant may reduce the participants' indifference. More emotionally provoking activities may even impress those who are not negatively affected as long as the reactions are not negatively directed against the LDI consultants themselves. From a purely business perspective, external consultants may therefore improve their likelihood of future business by delivering a combination of emotion-provoking challenges and intimidating contexts, using the employers' power position to frame the participant as the weak part. A combination of performance pressure, evaluation, and emotionally provoking interventions is not conducive to workplace learning (Todorova, Bear, & Weingart, 2014) but seems aligned with the employer's interests and avoids indifference in the participants. This would explain why our respondents most often express their reactions in terms of powerlessness and a loss of self-confidence, using adjectives such as “small,” “sidelined,” and “naked.”

Some of the more dramatic techniques applied by consultants may be a matter of fads. Looking back on the history of clinical psychology, it is obvious that the techniques stemming from group psychotherapy, therapeutic



communities, and gestalt therapy have had their golden ages among practitioners in leadership development (Roth & Fonagy, 1996). The value of these techniques has been shown to be of dubious value even in their original healthcare settings (Barlow, 2010; Roth & Fonagy, 1996; Wampold, 2001), and documentation of their effects in workplace environments does not seem to exist. The so-called T-groups and sensitivity groups in particular have been known as risky practices since the 1970s (Gottschalk, Pattison, & Schafer, 1971). The introduction of clinical psychological techniques in leadership development is interesting and even promising (Kets de Vries, 2005), though the differences in roles and organizational context may be ethically problematic (Berglas, 2002; Feldman & Lankau, 2005; Schein, 2006). It is possible that less intrusive and provoking techniques have given way to more solution-focused techniques within the wave of positive psychology with less pressure on emotional exposures. But self-reflection remains a relevant focus in leadership development (London, 2002) and it is by no means certain that the label “positive” or “solution focused” renders clinical techniques less risky in the hands of uncritical consultants (Ehrenreich, 2005, 2010). LDIs will always be intertwined with effects on workplace relationships and careers, and invitations to help relationships always raise issues of intimacy, vulnerability, and limitations of responsibility.

Nonetheless, unobtrusive techniques may be less likely candidates for systematic evaluation. Our data show that organizations use simple feedback surveys rather than thorough assessments of the actual learning outcomes. Companies that evaluate more seem to leapfrog onto climate surveys that seldom make specific links to the developmental process that must mediate the effect of Kirkpatrick’s second level. Neither do such surveys frequently question if this really matters, corresponding to Kirkpatrick’s fourth level. The focus seems to be on immediate emotional impacts of LDI activities, such as participant or employee satisfaction, to the neglect of more instrumental and cognitive dimensions in leadership and organizational development. Earlier evaluations of team building (Salas et al., 1999) and outdoor programs (Stoltz, 1992) have shown that LDIs mainly focus on emotional reactions there and then, with few transfer effects to the workplace. Updates on such studies show that there may be a trend toward more rigorous evaluations with improved outcomes (Klein et al., 2009). Taken together with the meta-analysis of Collins and Holton (2004), these studies indicate that organizations may overlook the need for programs to fit both organizations and participants. This is precisely what our findings seem to indicate. That is respondents who have experienced negative effects of LDIs describe that the activities are not suited to their needs and adapted to their situations, that they personally were not sufficiently prepared for the activities, and that external consultants were the main cause of their negative experiences.

We do not mean to say that emotionally challenging or even provoking experiences should be banned from leadership development, quite the contrary. Effective learning needs to be emotionally engaging, and challenging assignments are significant learning opportunities as long as they are balanced with perceived progress (“no pain no gain”) (Benjamin & O’Reilly, 2011; Todorova et al., 2014). Our data indicate that people embrace leadership development activities in general. Our respondents protested the gap that exists between the invested efforts and the perceived outcome. Judged by prevailing evaluation practices and the responses from our subjects, leadership development interventions seem more like acts of faith than professionally planned activities (Antonakis et al., 2011). Prerequisites such as job analysis, pedagogical requirements, and investment controls seem absent.

We need more evidence-based and professional approaches to leadership development in organizations.

As with other fads and fashions in management theory, practices that show the promise of giving a competitive edge gain the status of “must-have” ailments (Strang & Macy, 2001). In light of previous evaluation research, our data suggest that some organizations enter a dysfunctional relationship with external third-party providers. The client’s organization is ready to pay considerable sums of money for a process in which the professional requirements are left entirely to the external service provider. Eager to seem worth the expenditure, the external providers turn up the emotional heat but with questionable effects for the participants. In order to secure useful, beneficial effects from leadership development activities, organizations need to start using in-house competence to ensure sufficient contextual relevance (Axelrod & Cohen, 1999). They must then subject the processes to qualified evaluation routines that match the resources invested. All management practices may apparently backfire when applied as blind routines (Pfeffer & Sutton, 2006). Leadership development is not an exception.

### **Implications for practice**

We think our study documents how leadership development may not be unequivocally positive but implies risks



to both participants and organizers if not carried out properly. It also provides arguments for more evidence-based and professional approaches to leadership development in organizations. Handled incompetently, leadership development interventions may not only be a waste of time, money, and opportunities, but they may also actually be de-motivating or even outright dangerous to participants. We suggest that those responsible for LDIs in organizations should take the threat of fads and rogue suppliers more seriously than has previously been the case. We also recommend they be more persistent in both designing and evaluating leadership development programs.

### Limitations and suggestions for further research

Our study was merely explorative, based on convenience samples. Prospective evaluations of leadership development interventions from a broad range of organizations could link the processes of wasted and harmful LDI more explicitly to evaluation practices, spending, and consultant competence. Such evaluations may increase our knowledge about negative and even harmful effects of leadership development, which can be useful in preventing or managing the negative impact of leadership development interventions in the future. Our study is limited to Scandinavian organizations. Cultural and historical traditions may result in other findings elsewhere. Future research should therefore address how culture and professional development may affect both negative and positive outcomes of leadership development.

### References

- American Psychological Association (APA) (2010). *Ethical Principles of Psychologists and Code of Conduct*. Washington, D.C.: American Psychological Association.
- Antonakis, J., Fenley, M., & Liechti, S. (2011). Can charisma be taught? Tests of two interventions. *The Academy of Management Learning and Education*, 10(3), 374–396. doi:[10.5465/amle.2010.0012](https://doi.org/10.5465/amle.2010.0012)
- Aragón-Sánchez, A., Barba-Aragon, I., & Sanz-Valle, R. (2003). Effects of training on business results. *International Journal of Human Resource Management*, 14(6), 956–980. doi:[10.1080/0958519032000106164](https://doi.org/10.1080/0958519032000106164)
- Argyris, C. (2000). The relevance of actionable knowledge for breaking the code. In M. Beer & N. Nohria (Eds.), *Breaking the code of change* (pp. 429–445). Boston, MA: Harvard Business School Press.
- Arnulf, J. K. (2014). *A brief introduction to leadership*. Oslo: Universitetsforlaget.
- Arnulf, J. K., Larsen, K. R., Martinsen, Ø. L., & Bong, C. H. (2014). Predicting survey responses: How and why semantics shape survey statistics in organizational behavior. *PLoS ONE*, 9(9), 1–13. doi:[10.1371/journal.pone.0106361](https://doi.org/10.1371/journal.pone.0106361)
- Arthur, W. A., Bennett Jr., W., Edens Jr., P. S., & Bell, S. T. (2003). Effectiveness of training in organisations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(2), 234–245. doi:[10.1037/0021-9010.88.2.234](https://doi.org/10.1037/0021-9010.88.2.234)
- Avolio, B. C. (2004). *Leadership in balance: MADE/Born*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Avolio, B. J., Avey, J. B., & Quisenberry, D. (2010). Estimating return on leadership development investment. *The Leadership Quarterly*, 21(4), 633–644. doi:[10.1016/j.leafqua.2010.06.006](https://doi.org/10.1016/j.leafqua.2010.06.006)
- Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly*, 20(5), 764–784. doi:[10.1016/j.leafqua.2009.06.006](https://doi.org/10.1016/j.leafqua.2009.06.006)
- Axelrod, R. M., & Cohen, M. D. (1999). *Harnessing complexity: Organizational implications of a scientific frontier*. New York, NY: Free Press.
- Barlow, D. H. (2010). Negative effects from psychological treatments: A perspective. *American Psychologist*, 65(1), 13–20. doi:[10.1037/a0015643](https://doi.org/10.1037/a0015643)

- Baruch, Y., Bell, M. P., & Gray, D. (2005). Generalist and specialist graduate business degrees: Tangible and intangible value. *Journal of Vocational Behavior*, 67, 51–68. doi:[10.1016/j.jvb.2003.06.002](https://doi.org/10.1016/j.jvb.2003.06.002)
- Baruch, Y., & Leeming, A. (2001). The added value of MBA studies – graduates' perceptions. *Personnel Review*, 30(5–6), 589–601. doi:[10.1108/EUM00000000005941](https://doi.org/10.1108/EUM00000000005941)
- Benjamin, B., & O'Reilly, C. (2011). Becoming a leader: Early career challenges faced by MBA graduates. *Academy of Management Learning & Education*, 10(3), 452–472. doi:[10.5465/amle.2011.0002](https://doi.org/10.5465/amle.2011.0002)
- Berglas, S. (2002). The very real dangers of executive coaching. *Harvard Business Review*, 80(6), 86–92.
- Collins, D. (2000). *Management fads and buzzwords. Critical-practical perspectives*. London: Routledge.
- Collins, D. B., & Holton, E. F. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001. *Human Resource Development Quarterly*, 15(2), 217–248. doi:[10.1002/hrdq.1099](https://doi.org/10.1002/hrdq.1099)
- Cooper, C. (2010). *Individual differences and personality* (3rd ed.). London, UK: Hodder Education.
- Day, D. V. (2001). Leadership development: A review in context. *Leadership Quarterly*, 11(4), 581–613. doi:[10.1016/S1048-9843\(00\)00061-8](https://doi.org/10.1016/S1048-9843(00)00061-8)
- de Graaf, R., Tuithof, M., van Dorsselaer, S., & ten Have, M. (2012). Comparing the effects on work performance of mental and physical disorders. *Social Psychiatry and Psychiatric Epidemiology*, 47(11), 1873–1883. doi:[10.1007/s00127-012-0496-7](https://doi.org/10.1007/s00127-012-0496-7)
- DeNisi, A. S., & Kluger, A. N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of management executive*, 14(1), 129–139. doi:[10.5465/ame.2000.2909845](https://doi.org/10.5465/ame.2000.2909845)
- Dixon, N. (1976). *On the psychology of military incompetence*. London, UK: Basic Books.
- Ehrenreich, B. (2005). *Bait and switch: The (futile) pursuit of the American dream*. New York, NY: Metropolitan Books.
- Ehrenreich, B. (2010). *Smile or die: How positive thinking fooled America and the world*. London, UK: Granta.
- Feldman, D. C., & Lankau, M. J. (2005). Executive coaching: A review and agenda for future research. *Journal of Management*, 31(6), 829–828. doi:[10.1177/0149206305279599](https://doi.org/10.1177/0149206305279599)
- Fulmer, R. M., & Goldsmith, M. (2001). *The leadership investment: How the world's best organizations gain strategic advantage through leadership development*. New York, NY: American Management Association.
- Glaser, B. G., & Strauss, A. (1967). *Discovery of Grounded Theory. Strategies for Qualitative Research*. Mill Valley, CA: Sociology Press.
- Gottschalk, L. A., Pattison, E. M., & Schafer, D. W. (1971). Training groups, encounter groups, sensitivity groups and group psychotherapy. *Western Journal of Medicine (prev. California Medicine)*, 115(2), 87–93.
- Gurdjian, P., Halbeisen, T., & Lane, K. (2014). Why leadership development programs fail. *McKinsey Quarterly* (January). Retrieved from <http://www.mckinsey.com/global-themes/leadership/why-leadership-development-programs-fail>
- Hannum, K., Hannum, J., & Reinelt, C. (2007). *The handbook of leadership development evaluation*. Hoboken, NJ: John Wiley & Sons, Inc.
- Harms, P. D., Spain, S. M., & Hannah, S. T. (2011). Leadership development and the dark side of personality. *The Leadership Quarterly*, 22(3), 495–509. doi:[10.1016/j.leaqua.2011.04.007](https://doi.org/10.1016/j.leaqua.2011.04.007)
- Hedaa, L. (1997). *Sat ud af spillet: Case Tele Danmark Forlag – Personal Management Institute [Sidelined – the case of Tele Danmark Forlag and the Personal Management Institute]*. Frederiksberg, Denmark:

## Samfundslitteratur.

- Hill, A. (2011, November 25). Inside McKinsey. *Financial Times*. Retrieved from <http://www.ft.com/intl/cms/s/2/0d506e0e-1583-11e1-b9b8-00144feabdc0.html#axzz1g2apkxM7>
- Kets de Vries, M. (2005). Leadership group coaching in action: The Zen of creating high performance teams. *Academy of Management Executive*, 19(1), 61–76. doi:10.5465/ame.2005.15841953
- Kets de Vries, M., & Balazs, K. (2005). Organizations as optical illusions: A clinical perspective on organizational consultation. *Organizational Dynamics*, 34(1), 1–17 doi:10.1016/j.orgdyn.2004.11.006
- Khurana, R. (2007). *From higher aims to hired hands: The social transformation of American business schools and the unfulfilled promise of management as a profession*. Princeton, NJ.: Princeton University Press.
- Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training and Development Journal*, 33(6), 78–92.
- Klein, C., DiazGranados, D., Salas, E., Le, H., Burke, C. S., Lyons, R., & Goodwin, G. F. (2009). Does Team Building Work? *Small Group Research*, 40(2), 181–222. doi:10.1177/1046496408328821
- Kline, P. (2000). *Handbook of psychological testing* (2nd ed.). London, UK: Routledge.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. doi:10.1037/0033-2909.119.2.254
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Los Angeles, CA: Sage.
- Livingston, J. S. (1971). Myth of the well-educated manager. *Harvard Business Review* (Jan.–Feb.), 79–89. Retrieved from <https://hbr.org/1971/01/myth-of-the-well-educated-manager>
- London, M. (2002). *Leadership development: Paths to self-insight and professional growth*. Mahwah, NJ: Lawrence Erlbaum.
- Mabey, C., & Finch-Lees, T. (2008). *Management and leadership development*. London: Sage.
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293–334.
- McCauley, C. D., & Van Velsor, E. (Eds.). (2004). *The center for creative leadership: Handbook of leadership development* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Mintzberg, H. (2004). Leadership and management development: An afterword. *Academy of Management Executive*, 18(3), 140–142. doi:10.5465/AME.2004.14776190
- O'Leonard, K. (2010). *The corporate learning factbook: Statistics, benchmarks, and analysis of the U.S. corporate training market*. Bersin by Deloitte. Retrieved from <https://www.bersin.com/Store/Details.aspx?docid=103311678>
- Pfeffer, J., & Sutton, R. I. (2006). *Hard facts, dangerous half-truths, & total nonsense: Profiting from evidence-based management*. Boston, MA: Harvard Business School Press.
- Powell, K. S., & Yalcin, S. (2010). Managerial training effectiveness: A meta-analysis 1952–2002. *Personnel Review*, 39(2), 227–241. doi:10.1108/00483481011017435
- Quirk, M. P., & Fandt, P. M. (2000). *The 2nd language of leadership*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Roth, A., & Fonagy, P. (1996). *What works for whom? A critical review of psychotherapy research*. New York, NY: Guilford.
- Salas, E., Rozell, D., Mullen, B., & Driskell, J. E. (1999). The effect of team building on performance – An integration. *Small Group Research*, 30(3), 309–329. doi:[10.1177/104649649903000303](https://doi.org/10.1177/104649649903000303)
- Schein, E. H. (2006). Coaching and consultations revisited: Are they the same? In M. Goldsmith & L. Lyons (Eds.), *Coaching for leadership* (2nd ed., pp. 17–25). San Francisco, CA: Wiley.
- Stoltz, P. (1992). An examination of leadership development in the great outdoors. *Human Resource Development Quarterly*, 3(4), 357–372. doi:[10.1002/hrdq.3920030407](https://doi.org/10.1002/hrdq.3920030407)
- Strang, D., & Macy, M. W. (2001). In search of excellence: Fads, success stories, and adaptive emulation. *American Journal of Sociology*, 107(1), 147–182. doi:[10.1086/323039](https://doi.org/10.1086/323039)
- Takala, M., Winegar, D., & Kuusela, J. (2009). Leadership developmental needs – a system for identifying them. *Australian Journal of Adult Learning*, 49(1), 126–147.
- Taylor, P. J., Russ-Eft, D. F., & Taylor, H. (2009). Transfer of management training from alternative perspectives. *Journal of Applied Psychology*, 94(1), 104–121. doi:[10.1037/a0013006](https://doi.org/10.1037/a0013006)
- Taylor, S. N., & Bright, D. S. (2011). Open-mindedness and defensiveness in multisource feedback processes: A conceptual framework. *Journal of Applied Behavioral Science*, 47(4), 432–460. doi:[10.1177/0021886311408724](https://doi.org/10.1177/0021886311408724)
- Todorova, G., Bear, J. B., & Weingart, L. R. (2014). Can conflict be energizing? A study of task conflict, positive emotions, and job satisfaction. *Journal of Applied Psychology*, 99(3), 451–467. doi:[10.1037/a0035134](https://doi.org/10.1037/a0035134)
- Van Buren, M. E., & Erskin, W. (2002). *State of the industry: ASTD's annual review of trends in employer provided training in the United States*. American Society for Training & Development.
- Vardi, Y., & Wiener, Y. (1996). Misbehavior in organizations: A motivational framework. *Organization Science*, 7(1), 151–165. doi:[10.1287/orsc.7.2.151](https://doi.org/10.1287/orsc.7.2.151)
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ.: Lawrence Erlbaum Associates.
- Whetten, D., & Cameron, K. S. (2005). *Developing management skills*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Yukl, G. (2012). *Leadership in organizations (8th ed.)*. Harlow, UK: Pearson Education.

#### Citation

Arnulf, J. K., Glasø, L., Andreassen, A. K. B., & Martinsen, Ø. L. (2016). The dark side of leadership development: An exploration of the possible downsides of leadership development. *Scandinavian Psychologist*, 3, e18. <http://dx.doi.org/10.15714/scandpsychol.3.e18>

#### Abstract

### **The dark side of leadership development: An exploration of the possible downsides of leadership development**

In principle, leadership development may have positive effects, negative effects, or no effects at all. The present study aims to explore the potential negative effects of leadership development. We approach this issue with three studies. The first study develops theoretical reasons to expect negative effects and provides a qualitative description of such instances based on 14 semi-structured interviews with people witnessing negative effects.



The second study is a quantitative assessment of the prevalence of negative effects. While 97.4% of 189 managers responding to an anonymous survey have experienced positive effects from developmental activities, 63% of them have also experienced instances they deemed wasted or ineffective, and 37% have witnessed actual negative effects. A small but distinct group (13% of the entire sample and 36% of those witnessing negative effects) reported detrimental effects on mental health and private life. The third study explores the evaluation practices of 143 companies to examine why negative leadership development interventions may persist. The data indicates that negative effects co-exist with a lack of systematic evaluation practices. Our findings suggest that a combination of insufficient evaluation with outsourcing of leadership development activities may render organizations susceptible to fads. In turn, the chance of exposing participants to ineffective and even harmful experiences increases.

**Keywords:** consultancy, evaluation, harmful effects, leadership development.

**Author affiliations:** Jan Ketil Arnulf & Lars Glasø – BI Norwegian Business School, Oslo, Norway; Anette K. B. Andreassen – Itera, Oslo, Norway; Øyvind Lund Martinsen – BI Norwegian Business School, Oslo, Norway.

**Contact information:** [Jan Ketil Arnulf](#), BI Norwegian Business School, N-0442 Oslo, Norway.  
Email: [jan.k.arnulf@bi.no](mailto:jan.k.arnulf@bi.no).

**Received:** July 13, 2016. **Accepted:** October 31, 2016. **Published:** December 2, 2016.

**Language:** English.

**Competing interests:** The authors report no conflict of interest. The authors alone are responsible for the contents and writing of this paper.

**Acknowledgement:** The authors are grateful to Linn Maurud, MSc, and Kathrine Grasbekk, MSc, for their contributions to this project.

This is a peer-reviewed paper.

### **Jan Ketil Arnulf**

Jan Ketil Arnulf er professor ved [Handelshøyskolen BI](#). Han er spesialist i klinisk psykologi og arbeids- og organisasjonspsykologi. De siste årene har han konsentrert seg om hvordan ledelse utøves og skapes. Han er forfatter av boken [Hva er ledelse](#).

### **Lars Glasø**

Lars Glasø er professor i organisasjonspsykologi ved [Handelshøyskolen BI](#).

### **Anette K. B. Andreassen**

Annete K. B. Andreassen er utdannet innenfor markedskommunikasjon og jobber i kommunikasjons- og teknologiselskapet [Itera](#).

### **Øyvind Lund Martinsen**

Øyvind Lund Martinsen er professor ved Institutt for ledelse og organisasjon ved [Handelshøyskolen BI](#) i Oslo.