Article

The case against coaching

Erik de Haan

Unsurprisingly, strong and growing interest in the effectiveness of workplace and executive coaching is yielding an increasing and consistent body of significant findings. Firstly, over the past 25 years coaching has enjoyed sustained growth inside larger organisations, with processes, codes of conduct and qualifications becoming more and more standardised. This has helped researchers to increasingly find realistic setting for doing research. Secondly, coaching takes place in tightly contracted, delineated, one-to-one conversations which can be easily quantified for research purposes. With the interventions being limited to the conversations and the sessions normally taking place in a neutral venue, a natural laboratory situation for the measurement of effectiveness has emerged which cannot be found for adjacent fields such as mentoring, team coaching, process consultation, leadership development programmes and OD consulting. Over recent years, the focused study of 'adverse experiences' or 'negative side effects' of coaching has remained relatively small and mostly disjunct from the effectiveness studies. This article provides a full review and reappraisal of those studies into the case against coaching, integrating them with what is known about negative side effects within quantitative coaching research, and proposes a vision for carrying this research forward.
Keywords: *executive coaching; outcome research; effectiveness of coaching; side effects; null findings.*

Introduction

THE PAST 10 years have seen something like an explosion of coaching research, with many PhDs in the field and powerful comparison studies emerging. My recent summary of 160 original, rigorous, quantitative coaching studies (of which 38 were randomised controlled trials, with 11 within the last three years), plus a great many mentoring and health coaching articles (De Haan, 2021), shows that we have finally achieved a strong research base in coaching and a convincing case for the use of coaching as an instrument of organisational and leadership development.

In this contribution I would, however, like to look at the case that can be made against coaching, based on all the inconclusive and contrary findings in this same research literature, and also based on the case for negative side effects which has been made by a few researchers. It is good science to amplify that voice: give your opponents or the single 'counter example' the maximum amount of space and time. After all, their voices are the only chance that we might all be spurred on to greater levels of certainty: a single, convincing counter example can make a whole theoretical edifice crumble and perish, according to the lauded principle of 'falsifiability'.

Can we actually treat 'negative' outcomes separate from 'positive' ones?

Before summarising all the evidence we can muster against coaching, it is important to realise that the research on negative (side) effects of coaching presupposes that we all agree on what negative side effects are, or even that we could all agree for any outcome whether it is a positive or a negative one. But the world of coaching is not as simple as that. The assumption that we know what 'negative outcomes' are in an absolute sense can and should be questioned. After all, one

person's negative outcome can very well be another's palpable gain. If coaching works, then coaching can positively strengthen good as well as bad management, healthy as well as relatively toxic organisations. There may be 'very real dangers' (Berglas, 2002) in the use of coaching, even if so much of it has been shown to be beneficial for clients and their organisations.

Let us begin with a simple thought experiment – suppose a certain small dose of 'C' makes a great majority of people ('P') feel significantly happier. Then one would have demonstrated a 'happiness effect', at least for most of the 'C' users, and possibly also for some other people in their direct surroundings (or potentially, it could also have some adverse effects, say on their 'rivals' at work). Of course, there could be other adverse effects to such a potent elixir, for example, 'P' may become addicted to it, or their teeth might fall out, or their liver could stop functioning after 10 years of using 'C'. But even if this were not the case, if there are no adverse or side effects whatsoever, it is important to remember that 'P' would have to pay to obtain 'C' (and in some classic tales happiness cures have been notoriously expensive; remember for example Mephistopheles' cure where you have to pay with your soul). 'P' would agree to a cost in time or money or other expenses to obtain or engage with 'C'; these costs should always be listed under negative side effects for 'P' or for 'P's organisation that pays for 'C'.

Back to coaching. Arguably, any study that has not demonstrated a beneficial effect for coaching on the intervention group or others in the organisation has demonstrated a net negative effect of coaching. After all, coaching is costly both in terms of the fees that need to be paid to the highly qualified professionals and in terms of the time investment and the demands of the contract, which may include thinking about clear objectives, being open and vulnerable about your feelings and what you have tried already towards those objectives, and dealing with the challenges during the conversations, including any 'homework' or 'aporia' and other frustrations of not (yet) finding answers to your quest. Psychological costs could also include the loss of secondary gains or the deep pain, loss and uncertainty felt when consciously working through trauma and tackling or tearing down natural defences. Every single coaching study that does not find positive results will at the same time be a testament of these types of costs and could therefore be argued to have a negative net effectiveness result.

Qualitative studies into negative side effects of coaching and mentoring

The qualitative study of negative side effects started earlier in the mentoring literature, where Eby et al. (2000) studied negative mentoring experiences for 277 mentees and Eby and McManus (2004) did the same for 90 mentors. Eby et al. (2008) were able to collect perspectives from both mentors and mentees in 80 matched mentor-mentee pairs in two large universities. They found that mentors' perceptions of negative mentoring experiences were related to both mentors' and mentees' perceptions of relationship quality and fair exchange. Burk and Eby (2010) also investigated the consequences of these negative side effects on mentees, their decisions to leave the mentoring relationship and what stops them leaving, for example, lack of perceived mentoring alternatives and fear of mentor retaliation.

Schermuly and others similarly started out with three explorative, qualitative studies, investigating

- 104 coaches' perceptions of negative effects of coaching for themselves in their most recent assignment (Schermuly, 2014), where the coaches reported the usual effectiveness ratings for their own coaching but also that 90 per cent of cases had negative side effects for them.
- 123 coaches' perceptions of negative effects of coaching for their most recent client (Schermuly et al., 2014), with 57 per cent coaches reporting negative side effects for their coachees.

- 111 clients' perceptions of negative effects of coaching for themselves in a recent assignment (Graßmann & Schermuly, 2016), with 68 per cent of clients reporting negative side effects.

Further measurements over a time period of eight weeks based on self-scores only and no controls, showed that

1. For clients the number of negative effects was inversely proportional to relationship quality at both measurement times and also to the coach's expertise at Time 1 (Graßmann & Schermuly, 2016); and
2. For coaches the number of negative effects was inversely proportional to the relationship quality and to their own feelings of competence as a coach, and also led to them perceiving more stress and impaired sleep eight weeks later (Graßmann et al., 2019).

Graßmann and Schermuly (2018) conducted an experiment with 29 coach-client dyads (Master's students who were coaching bachelor's students) of which a randomised 16 coaches received two two-hour group supervision sessions during the period of coaching. After the first coaching session, coaches took a questionnaire on neuroticism. After coaching completion, coaches and clients evaluated negative effects of coaching for clients. There was no correlation between these two estimates (contrary to Eby et al., 2008, in the mentoring domain). Coaches additionally evaluated negative effects of coaching for themselves, and these turned out to strongly predict the negative effects they saw for their clients. Coaches' neuroticism also predicted the amount of negative side effects they themselves saw but not the number of negative side effects that their clients experienced. The presence or absence of supervision made no difference to the amount of negative side effects reported.

Overview of what we know quantitatively about side effects

Let us first examine a few studies that did find adverse effects or negative side effects (which of the two it is depends on what you would choose to see as the main effect of coaching, i.e. the distinction is slightly arbitrary).

The worst news for coaching that I have been able to find were five (clusters of) studies where an adverse effect was found instead of the anticipated beneficial effect:

1. In an excellent independent study of a single consulting intervention, Aust et al. (2010) show how not to implement organisation-development in a hospital. Many consulting interventions, including leadership coaching, failed with better results recorded for the no-intervention control group participants. It appears that firstly, hospital employees did not have time to organise and conduct group meetings and, despite the fact that consultants repeatedly offered their help, employees only contacted them twice during the entire project; and secondly, leaders were unsure about their role in the intervention project and therefore never fully took ownership for it.
2. Bozer et al. (2013) showed that in their experiment the control group performed better in terms of line-manager-rated task performance than the intervention group, which had no less than 10–12 coaching sessions. It is important to note, however, that (a) the control group was not random: they were recruited as peers of the clients, and (b) these peers had a much lower (manager-rated) performance initially, after which (c) the peers' performance scores only climbed up to the same level as that of the clients, which had not moved significantly despite the 10 coaching sessions. With a much smaller non-randomised control group one can never be sure whether or not the 'Hawthorne' effect in that group outweighed the coaching effect in the other group.
3. Ebner et al. (2020) also find a negative impact of the single one-to-one coaching session that they offered – it significantly

reduced the client's life satisfaction. That seems quite a powerful result of a single session, and contrary to the direction expected. Moreover, it is interesting to note that 'decrease in life satisfaction' is one of the top client negative side effects that Graßmann and Schermuly (2018) found – a few places below the most frequent one, 'decrease in job satisfaction'. Again, group assignment was not random in a similar way as in Bozer et al. (2013): participants chose which experimental group they wanted to belong to. Moreover, Ebner et al. (2020) tested only a single coaching session, so the lack of satisfaction could have come from scarcity of the offer.

4. Building on Wageman's (2001) result that coaching does not help much if a proper team structure is not in place, Carson et al. (2007) and Buljac-Samardzic and Van Woerkom (2015) show that it is possible to coach too much in some teams. They showed that, although it was possible to demonstrate a correlation between (managerial) team coaching now and team effectiveness a year later, this relationship was only significant when team reflection (Buljac-Samardzic & Van Woerkom, 2015) or the team's supportive environment (Carson et al., 2007) had been low. There was indeed a detrimental effect on efficiency when team reflection was high: something they labelled 'excessive managerial coaching'. It is perhaps understandable that there is a case for 'excessive' reflection and coaching, for example, at those moments where action is needed. An executive coach who does not pick up the fact that a client is bringing something that s/he has to act on immediately and is looking for an immediate outcome but carries on inviting the client to engage in deep and abstract reflection, could be a similar and understandable example of ineffective coaching. At the very least, we should as executive coaches make it explicit that we notice the urgency, and contract around an immediate deliverable if that is what our client wants from the session.
5. Another interesting finding was that coaching needs the addition of instruction to make a difference on objective academic marks (Franklin & Doran, 2009) and otherwise does not do very much beyond increasing self-scores. Hui et al. (2013) and Zanchetta et al. (2020) also explicitly show that in some cases guidance and training are better suited for certain outcomes, such as cognitive acquisition. However, they did also find that facilitative coaching had better results for new tasks. One can understand this as a particular application for coaching (towards new strategies and more creative tasks) but also as a negative result, namely when it comes to academic achievement, cognitive acquisition and repeat tasks. Similarly, Deane et al. (2014) found that 'transformational' coaching had a more lasting impact on productivity than 'skills' coaching had.

I would argue that none of these studies demonstrates a strong case for an adverse effect for the use of executive coaching: for the first one, the overall intervention seems to be doomed from the beginning (and in fact one can argue that precisely the coaching element for the leaders was more successful than the rest of the project, because it was entirely voluntary and still all leaders took up the offer and took seven hours of coaching on average); for the next two studies, one doubts the significance, especially given the fact that there are many other studies which do show a positive effect on the same outcome dimensions (De Haan, 2021); and the latter two groups of studies list very specific circumstances where coaching might not work or not work so well – namely in those circumstances where reflection is already high and there are perhaps other expectations of the helper, e.g. to help move the team to action (Buljac-Samardzic & Van Woerkom, 2015) or for the coachee to learn more facts (Hui et al., 2013, and Zanchetta et al., 2020).

Now let us look at the studies that failed to demonstrate any effects (beyond self-scored) where positive change had been expected. I will skip the (pilot) studies with 'null' effects that were so small in scale that one would not necessarily have expected much significance in terms of effectiveness, such as Miller (1990) and Tee et al. (2017).

There are several studies, starting from the early Green et al. (2006, 2007) and Grant et al. (2009, 2010) studies, where a significant effect was demonstrated on client scores but not on peer ratings. The same was found again on coach scores (Schermuly et al., 2020), direct-report scores (Finn et al., 2007), manager scores (Williams & Lowman, 2018; Jones et al., 2019), peer, direct-report and manager scores (Nieminen et al., 2013), and some objective measures (e.g. physiological responses: Schermuly et al., 2020, and Howard, 2015; and business viability: Oberschachtsiek & Scioch, 2015). These studies do stand apart as a warning to coaches and researchers alike. They show that demonstrating universally agreed, objective results from coaching is still a difficult job, and that several who have set out to do so have failed (they can only report significant findings on self-scores, which we know are notoriously biased towards false positives).

Oberschachtsiek and Scioch (2015) studied a large-scale (N>418.000), historic dataset of German support structures for new business owners over five years, which should be very relevant for commissioners and decision makers in the field of executive coaching, as it makes a good link between expense and return on investment in coaching. They focused on three different coaching programmes provided along with a financial subsidy to an entrepreneur who started a business while they were unemployed. Their results show that the coaching effects tended to be low and did not make a meaningfully significant difference in the viability of their businesses (they are only significant because of the very large sample size). On the basis of timings in their data of foreclosing businesses, they can argue that some of the coaching is likely to have led to better insight about the viability of the business and, therefore, despite the businesses of these coached entrepreneurs floundering, some entrepreneurs may have gotten out earlier thanks to the coaching.

Oberschachtsiek and Scioch (2015) seems a very rigorous study and, although entirely historical, it is perhaps one of the strongest arguments we have against coaching, as this work clearly makes it difficult to argue for coaching as part of the subsidies for new entrepreneurs. However, the authors argue that possibly coaching did help the entrepreneurs in many cases, namely by clarifying for them when to bow out of their fledgling enterprises, before giving up became too late and too costly.

What does it mean for coaching practice

In the previous section we have found some clear results on negative effects in coaching. Sometimes it has not been possible to prove a positive objective result, such as on (reduction of) stress hormones or the viability of businesses. Sometimes it is not possible to show positive results on peer, line-manager and direct-report ratings. And sometimes even self-scores do not show significance between intervention and control group. It is worth mentioning that there are now twelve studies in which such null results have been found, as opposed to the 148 other quantitative studies in de Haan (2021) that were mostly more rigorous (e.g. many had larger sample sizes and randomsed control groups) and did show significant effects in workplace and executive coaching. As mentioned in the previous section, a few of those 'negative' studies showed that only specific applications of coaching work towards the intended target, that is, they seem to say that sometimes there is no need for still more reflection or that directive 'skills' coaching has its limits when it comes to dealing with new challenges and the need for creativity.

It seems that the case against coaching remains weak for the time being, with mostly evidence for clear circumstances where

coaching might not add much or even be 'too much', that is, circumstances where more reflection is not required. The case against coaching has nevertheless one strong argument in its favour. It is evidenced that in social sciences and psychology only around 20 per cent of 'null results' actually get published. This is not so much due to journals not being interested in null results, but more due to stubbornness of researchers who often stick with their models against evidence and tend to conclude that their 'experiment did not work out' and is therefore not worth publishing (see Franco et al., 2014). This means that, if we have 12 negative and null results, this is more likely to signify 60 independent studies that showed a null result, of which 48 were not published. Moreover, the other 147 articles in De Haan (2021) do not all demonstrate coaching effectiveness, as many of those studies only compare conditions. If I take those latter studies out, only 93 studies do provide evidence of effectiveness, which means that the odds would be around two to three to find a null result.

In sum, on the one hand focusing on the negatives in coaching research has given us more confidence for the effectiveness of coaching, i.e. only a stronger case for executive and workplace coaching. On the other hand, the analysis of potential adverse effects in coaching remains important and needs to be taken up with the help of better definitions of what constitutes positive and negative impact. For the time being, we can conclude that 87 per cent of the studies published to date show at least a few significantly positive outcomes.

About the author

Erik de Haan studied Theoretical Physics and undertook his PhD in Psychophysics and MA in Psychodynamic Psychotherapy. He is Director of the Ashridge Centre for Coaching at Hult International Business School, UK, and Professor of Organisation Development at the VU University in Amsterdam, the Netherlands. He is the programme leader of Ashridge's MSc in Executive Coaching and PG Diploma in Organisational Supervision. He has published more than 200 professional and research articles and 14 books, covering his expertise as an organisational consultant, therapist, and executive coach.

Correspondence

Erik de Haan
erik.dehaan@ashridge.hult.edu

References

Aust, B., Rugulies, R., Finken, A. & Jensen, C. (2010). When workplace interventions lead to negative effects: learning from failures. *Scandinavian journal of public health, 38(3_suppl)*, 106–119.

Bozer, G., Sarros, J.C. & Santora, J.C. (2013). The role of coachee characteristics in executive coaching for effective sustainability. *Journal of Management Development, 32*(3), 277–294.

Buljac-Samardzic, M. & Van Woerkom, M. (2015). Can managers coach their teams too much? *Journal of managerial psychology, 30*(3), 280–296.

Burk, H.G. & Eby, L.T. (2010). What keeps people in mentoring relationships when bad things happen? A field study from the protégé's perspective. *Journal of Vocational Behavior, 77*(3), 437–446.

Carson, J.B., Tesluk, P.E. & Marrone, J.A. (2007). Shared leadership in teams: An investigation of antecedent conditions and performance. *Academy of Management Journal, 50,* 1217-1234.

De Haan, E. (2021). *What works in executive coaching: Understanding outcomes through quantitative research and practice-based evidence.* London & New York: Routledge.

Deane, F.P., Andresen, R., Crowe, T.P. et al. (2014). A comparison of two coaching approaches to enhance implementation of a recovery-oriented service model. *Administration and Policy in Mental Health and Mental Health Services Research, 41*(5), 660–667.

Ebner, K., Souček, R. & Kauffeld, S. (2020). Incongruities between values, motives, and skills: Exploring negative effects of self-exploration in career coaching. *British Journal of Guidance & Counselling,* 1–23.

Eby, L.T. & McManus, S.E. (2004). The protégé's role in negative mentoring experiences. *Journal of Vocational behavior, 65*(2), 255–275.

Eby, L.T., Durley, J.R., Evans, S.C. & Ragins, B.R. (2008). Mentors' perceptions of negative mentoring experiences: Scale development and nomological validation. *Journal of Applied Psychology, 93,* 358–373.

Eby, L.T., McManus, S.E., Simon, S.A. & Russell, J.E. (2000). The protege's perspective regarding negative mentoring experiences: The development of a taxonomy. *Journal of Vocational Behavior, 57*(1), 1–21.

Finn, F.A., Mason, C.M. & Bradley, L.M. (2007). *Doing well with executive coaching: Psychological and behavioural impacts.* Paper presented at the Academy of Management Annual Meeting Proceedings, Philadelphia, PA.

Franco, A., Malhotra, N. & Simonovits, G. (2014). Publication bias in the social sciences: unlocking the file drawer. *Science, 345,* 1502–1505.

Franklin, J. & Doran, J. (2009). Does all coaching enhance objective performance independently evaluated by blind assessors? The importance of the coaching model and content. *International Coaching Psychology Review, 4*(2), 128–144

Grant, A.M., Curtayne, L. & Burton, G. (2009). Executive coaching enhances goal attainment, resilience and workplace wellbeing: A randomised controlled study. *The Journal of Positive Psychology, 4*(5), 396–407.

Grant, A.M., Green, L.S. & Rynsaardt, J. (2010). Developmental coaching for high school teachers: Executive coaching goes to school. *Consulting Psychology Journal: Practice and Research, 62*(3), 151–168.

Graßmann, C. & Schermuly, C.C. (2016). Side effects of business coaching and their predictors from the coachees' perspective. *Journal of Personnel Psychology, 15,* 152–163.

Graßmann, C. & Schermuly, C.C. (2018). The role of neuroticism and supervision in the relationship between negative effects for clients and novice coaches. *Coaching: An International Journal of Theory, Research and Practice, 11*(1), 74–88.

Graßmann, C., Schermuly, C.C. & Wach, D. (2019). Potential antecedents and consequences of negative effects for coaches. *Coaching: An International Journal of Theory, Research and Practice, 12*(1), 67–88.

Green, L.S., Grant, A.M. & Rynsaardt, J. (2007). Evidence-based life coaching for senior high school students: Building hardiness and hope. *International Coaching Psychology Review, 2*(1), 24–32.

Green, L.S., Oades, L.G. & Grant, A.M. (2006). Cognitive-behavioural, solution-focused life coaching: Enhancing goal striving, wellbeing, and hope. *The Journal of Positive Psychology, 1*(3), 142–149.

Howard, A.R. (2015). Coaching to vision versus coaching to improvement needs: A preliminary investigation on the differential impacts of fostering positive and negative emotion during real time executive coaching sessions. *Frontiers in psychology 6,* 455.

Hui, T.Y., Sue-Chan, C. & Wood, R.E. (2013). The contrasting effects of coaching style on task performance: The mediating roles of subjective task complexity and self-set goal. *Human Resource Development Quarterly, 24*(4), 429–429.

Jones, R.J., Woods, S.A. & Zhou, Y. (2019) The effects of coachee personality and goal orientation on performance improvement following coaching: A controlled field experiment. *Applied Psychology.* doi/abs/10.1111/apps.12218.

Miller, D.J. (1990). The effect of managerial coaching on transfer of training. *Dissertation Abstracts International Section B, 50* (2435).

Nieminen, L.G., Smerek, R., Kotrba, L. & Denison, D. (2013). What does an executive coaching intervention add beyond facilitated multisource feedback? Effects on leader self-ratings and perceived effectiveness. *Human Resource Development Quarterly, 24*(2), 145–176.

Oberschachtsiek, D. & Scioch, P. (2015). The outcome of coaching and training for self-employment. A statistical evaluation of outside assistance support programmes for unemployed business founders in Germany. *Journal for Labour Market Research, 48,* 1–25.

Schermuly, C.C. (2014). Negative effects of coaching for coaches: An explorative study. *International Coaching Psychology Review, 9,* 167–182.

Schermuly, C.C., Schermuly-Haupt, M.L., Schölmerich, F. & Rauterberg, H. (2014). Zu Risiken und Nebenwirkungen lesen Sie – Negative Effekte von Coaching (For risks and side effects read – negative effects of coaching). *Zeitschrift für Arbeits- und Organisationspsychologie, 58,* 17–33.

Schermuly, C.C., Wach, D., Kirschbaum, C. & Wegge, J. (2020). Coaching of insolvent entrepreneurs and the change in coping resources, health, and cognitive performance. *Applied Psychology (In Press).*

Tee, D., Shearer, D. & Roderique-Davies, G. (2017). The client as active ingredient: 'Core self-evaluations' as predictors of coaching outcome variance. *International Coaching Psychology Review, 12*(2), 125–132.

Wageman, R. (2001). How leaders foster self-managing team effectiveness: Design choices versus hands-on coaching. *Organization Science, 12,* 559–577.

Williams, J.S. & Lowman, R L. (2018). The efficacy of executive coaching: An empirical investigation of two approaches using random assignment and a switching-replications design. *Consulting Psychology Journal: Practice and Research, 70*(3), 227–249.

Zanchetta, M., Junker, S., Wolf, A.M. & Traut-Mattausch, E. (2020). 'Overcoming the fear that haunts your success' – The effectiveness of interventions for reducing the impostor phenomenon. *Frontiers in Psychology 11,* 405.