Information, Attention and Training: Why Consulting Matters in Human Capital Intensive Organizations.

Evidence from a Field Experiment on Teacher Coaching

Simon Calmar Andersen
sca@ps.au.dk

Bastien Michel
bmichel@econ.au.dk

Helena Skyt Nielsen
hnielsen@econ.au.dk

Aarhus University

Trygfonden's Centre for Child Research

September 13th, 2017

1. Introduction

Recent studies have demonstrated substantial effects of consultancy on learning and productivity (Allen et al., 2011; Bloom et al., 2013; Hanna et al., 2014). However, the channels through which the effect of consultancy materializes and the relative strength of the different channels remain to be clarified. Studying the effect of consultancy on Indian textile firms, Bloom et al. (2013) found that the main effect of consultants was to deliver *information*. On the contrary, in their study of Indonesian seaweed farmers, Hanna et al. (2014) suggested that consultants' main effect is to draw *attention* to key input factors that are known, but not noticed by. While both of these channels, information and attention may be very relevant to various settings, we suggest that consultancy may have an effect through a third channel, which may be particularly important for human capital-intensive organizations (such as schools, research, etc.): *training*.

Our study investigates the impact of consultancy in the form of a one-year coaching program, which aims to improve teacher skills through peer coaching. In order to measure the impact of the program, we will implement a randomized controlled trial as part of which a randomly generated subset of teachers will be selected to benefit from the intervention. We take advantage of this experiment to explore the strength of the following three mechanisms through which coaching can improve teaching skills: information, attention, and training. While the importance of information, attention, and training have been studied separately in previous studies, this is to our knowledge the first time one tries to disentangle their relative strength in an experimental setting. Especially, we examine the hypothesis that coaching can have a separate effect on capabilities (through training) in human capital organizations – besides any effects on knowledge (based on new information) and noticing (based on drawing attention to input factors).

In carrying out this research project, we aim to provide additional evidence on the impact of co-teaching programs and, more generally, understand the mechanisms driving the results. In turn, this

would allow us to understand why human capital-intensive organizations do not reach their production-possibility frontier and what can be done to make them more efficient.

2. Theory

Just as for firms' production process, children's knowledge acquisition process can be represented with the help of a production function (the education production function), taking the key factors influencing children's learning process as input factors (Todd and Wolpin, 2003). Naturally, many of the inputs are directly related to the school environment, where a large part of children's learning take place, and to their teachers' competencies. Unfortunately, evidence suggest that there is a wide variation in teachers' competencies, which in turn results in important inefficiencies in the production of knowledge and inequalities across children.

Coaching can help teachers get closer to the production-possibility frontier through at least two different channels. Indeed, coaching can first provide teachers with information on the education production function (e.g. its functional form, its input factors, etc.) thereby increasing their knowledge (Bloom et al., 2013). Second, coaching can help draw teachers' attention to key input factors they were previously not noticing (Hanna et al., 2014; Schwartzstein, 2014). This can happen because teachers may already have priors with respect to what the key input factors are, but also because the education production function takes many factors as inputs and teachers can only pay attention to a limited number of them. As a consequence, while learning can occur when a teacher pays attention to the extent to which they use a given input (by using that knowledge to assess the impact of that input factor on children's learning), it does not occur when a teacher does not pay attention to a given input: the extent to which the input factor is used becomes random and hard to observe *a posteriori*, which in turn prevents any kind of inference. In this context, teachers may be brought closer to the production-possibility frontier by fostering learning through noticing. This may be achieved in several ways and, in particular by encouraging teachers to pay attention to the level of some specific inputs and/or by revealing information pertaining to the impact of these input factors on the output level.

However, we anticipate that the most important channel through which coaching has an impact on children's learning process is through teacher training.[1] Indeed, manipulating the level of certain inputs may be harder for teachers than for farmers (as described in Hanna et al. (2014)) or managers (as described in Bloom et al. (2013)). Indeed, while Indonesian seaweed farmers can easily adjust sequentially pod size and pod interval, and Indian managers can easily implement sequentially uncluttered factory floors and safe environments, teachers often have to attend to different harder-to-change input factors at the same time (using visual support, modulating voice tone, exploiting body

[1] In particular, we expect that learning through noticing can be more difficult in the field of education because it implies being able to observe output levels, which is often more difficult in this context than in the examples described in Hanna et al. (2014) and Bloom et al. (2013). In these examples, output measures are quickly available to the farmer and the manager. Instead, test scores measure a fraction of skills accumulated and, in some cases, such measures are only available with after a long period of time.

language, interacting with students, etc.). The complexity of this multi-dimensional task suggests that teachers might benefit from supervised repeated practice to improve on their performance. Eventually, regular practice and feedback from a coach may change certain teaching techniques from requiring effortful attention (Kahneman, 1973; Hanna et al., 2014) until they become automated and intuitive – part of what Kahneman describes as brain's system 1 decision making process (Kahneman, 2011).

3. Intervention

As part of this study, we aim to measure the impact of a teacher-coaching intervention. We will do so in a context where all teachers (whether or not they receive the coaching intervention) will benefit from a professional development course designed to improve their teaching skills. Half of those teachers will be selected to benefit from a coaching intervention as well.

3.1 Professional development course

We will implement an intervention where all participating teachers and coaches attend five one-day course modules providing them with *information* regarding the following education production factors: i) Physical surroundings, ii) Structuring schoolwork, iii) Self-regulation iv) The teacher's focus, and v) The teacher's communication.[2] The course modules will draw teachers' *attention* to five groups of inputs (teaching techniques) that facilitates teaching in an inclusive classroom setting.[3] As part of these course modules, teachers will learn to use both technological input factors (e.g. timer, checklist, five-point scale, etc.) and behavioral input factors (e.g. support in transition phases, use calm voice, mentalizing, etc.). Furthermore, during each module, teachers will be asked to reflect on their own students and to set an agenda for themselves detailing how they intend to change their teaching practices so as to take into account what they will have learnt during the course modules. In addition, each course module will start out with rehearsal and reflection on what they have learnt in order to consolidate knowledge.

In doing so, control and treated teachers will be equipped with the same level of *information* in relation to effective teaching techniques and will be encouraged to pay the same level of *attention* to relevant input factors.

The course modules will be held on the following dates:

- Course module 1: August 9th, 2017
- Course module 2: September 14th, 2017
- Course module 3: October 31st, 2017
- Course module 4: November 28th, 2017

[2] In Danish: i) Fysiske rammer, ii) Rammer for skoleopgaver, iii) Regulering af sanser, opmærksomhed og følelser iv) Den voksnes fokus og v) Den voksnes kommunikation.

[3] Our intervention is inspired by the ASD NEST support project http://steinhardt.nyu.edu/asdnest/, which is designed to include students with autism-spectrum disorders in ordinary age-appropriate classrooms.

- Course module 5: February 7th, 2017

3.2 Treatment: Teacher-coaching program

Teachers selected to receive the teacher-coaching program will be paired with a co-teacher, their "coach", with whom they will spend 2.5 hours per week during an entire school year. Co-teachers will help teachers get closer to the production-possibility frontier through "peer coaching", as part of which the teacher receives feedback on how to improve teaching practice.[4] These sessions will be dedicated to the implementation of what will have been covered during the professional development courses.

As already mentioned earlier, coaching (or consultancy) may theoretically affect teachers through at least three channels: information, attention drawing, and training. We expect that most of the impact of the co-teaching intervention will arise because of a training effect. We further expect that the intervention's impact will be greater on *behavioral* input factors (e.g. use direction instead of correction, support in transition phases, use calm voice, etc.) than on *technological* input factors (e.g. timer, check list, etc.), because behavioral input factors require more training in order to change.

4. Research design

4.1 Study population

We recruited teachers among all public schools located in Aarhus municipality, the second largest municipality in Denmark, inviting all language arts in grade levels 3-6 to participate. In total, 17 schools signed up to participate in the experiment with an average of four language arts teachers per school – representing 68 language arts teachers in total.

Mathematics teachers were also invited to sign up for the intervention. Unfortunately, relative financial constraints (combined with a more limited number of mathematics teachers who actually signed up for the intervention) led us to leave them out of the coaching experiment. However, they will be given the opportunity to follow the Professional Development Course.

Each school will appoint one or more persons to be their school's coach(es). The appointed person(s) will be requested to have received a specific training on how to handle students with special needs (learning difficulties; social, emotional, or behavioral problems).

4.2 Trial design

Randomization

As there are strong reasons to expect significant selection with respect to the type of teachers who enroll in a coaching program, we will measure its impact using a randomized controlled trial design.

[4] In the teaching material, the pairs are invited to use three different co-teaching models (one teaches, one observes; one teaches, one assists; co-teaching on equal terms).

The coaching intervention will be randomized at the teacher-level. For implementation-related reasons, it was decided that the draw would be stratified by school. As a consequence, we will randomly select in each school half of the language arts teachers to benefit from the coaching intervention. In schools where an uneven number of teachers will be enrolled in the experiment, the number of teachers who should benefit from the coaching intervention was decided in advance for logistical reasons. We will adjust for the ensuing variation in teachers' probability to receive the coaching intervention across schools in our estimation strategy.

When schools will have appointed more than one coach, the teacher-coach pairs will be randomly generated.

In total, 35 language arts teachers will receive the coaching intervention. Teachers and coaches will be informed of the results of the draw between the first and second course module (after completion of the baseline data collection).

Spillovers

The within-school randomization comes at the cost of potential spillovers across treatment and control teachers. We will assess the importance of these spillover effects by investigating if any impact of the co-teaching intervention can be found on control teachers and non-treated students[5] when the share or the number of language arts teachers teaching at the same grade level increases (variation which, in our setting, should be random).

Effect of the professional development course

In order to measure the effect of the professional development course itself (capturing the impact of increasing teachers' information and attention – holding the amount of training they receive constant), we will compare our sample of *control* pupils with i) pupils enrolled in non-participant schools in Aarhus Municipality and ii) pupils enrolled in schools located in Denmark but outside of Aarhus Municipality. Under the assumption that pupil outcomes would have followed parallels trends in *control* schools and in the two comparison groups i) and ii), we should be able to measure the causal impact of the professional development course using a Difference-in-Difference approach.

4.3 Research questions

The trial is designed to study the following primary research question:

- **Q1:** What is the impact of the coaching intervention on teaching practices?

As part of this research project, we also aim to answer the following secondary research questions:

- **Q2:** What is the relative strength of the information, attention, and training mechanisms?
- **Q3:** What is the impact of the coaching intervention on students' achievements?

[5] Students enrolled in control classes and, whenever possible, students enrolled in classes left out of the experiment.

- **Q4:** What are the characteristics of the students most likely to benefit from the intervention? In particular, we would be interested in assessing whether or not the effect varies with respect to students' cognitive abilities and gender.
- **Q5:** What teaching skills – technological or behavioral – are most likely to benefit from the intervention?

4.4 Data sources and outcome variables

Rounds of data collection and data sources

In order to answer our research questions, several rounds of data collection will be carried out (for each batch) at different points in time:

1/ At the beginning of the school year, during the first 20 minutes of the first course module (*baseline data collection,* August 2017 – before the results of the randomization are shared with the schools)
2/ During the first 20 minutes of each remaining course module (*follow-up surveys*)
3/ After the first coaching session carried out after each course module (*monitoring data collection*)
4/ At the end of the school year, upon completion of the course modules (*endline data collection,* June 2018)

The data will be gathered through different data sources:

- Administrative registers: student, parent and teacher background, tracking of students and teachers, output measures (e.g. national test scores and grades for the teacher's past, current, and future students)
- Observation data: classroom instrument developed from the course material and/or CLASS-UE measurement
- Teachers and students surveys: they will be collected online using the SurveyXact platform

Monitoring data

We will collect information depicting how the intervention is carried out. In particular, we will collect information on the identity of the teachers attending the professional development course and on those actually receiving the coaching intervention.

Furthermore, descriptive information will also be collected so as to understand how the coaching intervention is implemented. Teachers will be asked to answer (together with their coach) a short questionnaire after the first coaching session carried out after each course module. This questionnaire will collect information on their experience with the co-teaching intervention.

Outcome variables

Our primary outcome is teachers' teaching skills and, more specifically, whether or not they use the elements taught during the professional development course in class. This will be assessed by surveyors who will attend teachers' class and will rate the extent to which they use each of the

techniques taught during the course. We will supplement these class observations by a measure of teaching skills as reported by the students using a validated index of teaching quality (TRIPOD).

Our secondary outcome is students' achievements and, in particular, their reading skills (as measured by national standardized reading tests) and socio-emotional skills (as measured by national standardized wellbeing surveys).

Data will also be collected in subsequent years to measure two distinct types of long-term impacts. First, we will collect data on the subsequent cohorts taught by the teachers enrolled in the experiment. This will allow us to investigate if the coaching sessions has any long-term impact on their teaching practices or if its effect fades-out when coaching sessions stop. Second, we will track the students enrolled in the study in order to investigate whether or not control students catch up with time or if, on the contrary, gaps remain or even widen.

Measure of potential mechanisms

Knowledge. A first channel through which we can expect coaching to have an impact on teachers is through an increase in teachers' level of *knowledge* related to effective teaching techniques. In order to investigate the strength of this mechanism, teachers will be asked to answer a questionnaire investigating their perceived importance of various input factors. Teachers will be requested to answer these questions at baseline, endline, and at the beginning of each of the four remaining course modules.

Falling in with Hanna et al. (2014)'s methodology, we will proxy whether or not a teacher is knowledgeable about a given input factor by whether or not she is capable of assessing what its optimal level is. More precisely, a teacher will be considered knowledgeable about a given input factor if she does not answer "I don't know" to the question investigating what teachers think is this input factor's optimal level.

Given that treatment and control groups will benefit from the exact same course modules, we expect *information* to rise in similar ways in both groups immediately after the sessions. Nevertheless, differences might be found over time across groups and, if this were to happen, they would have to be attributed to the coaches' actions as per our experimental design.

Attention. A second channel through which we can expect coaching to have an impact on teachers is through an increase in teachers' level of *attention* to relevant input factors. In order to investigate the strength of this mechanism, teachers will be asked to answer a questionnaire investigating the extent to which they use these different input factors when teaching. Again, teachers will be requested to answer these questions at baseline, endline, and at the beginning of each of the four remaining course modules.

Falling in with Hanna et al. (2014)'s methodology, we will proxy whether or not a teacher pays attention to a given input factor by whether or not she is capable of assessing the extent she uses that input factor when teaching. More precisely, a teacher will be considered as paying attention to an input factor if she does not answer "I don't know" to the question investigating the extent to which teachers use that input factor.

Given that treatment and control groups will benefit from the exact same course modules, we expect *attention* to rise in similar ways in both groups immediately after the sessions. Nevertheless, differences might be found over time across groups and, if this were to happen, they would have to be attributed to the coaches' actions as per our experimental design.

Training. A third and final channel through which we expect coaching to have an impact on teachers is through *training*. Indeed, as described above, the specificities of the teaching job is such that they may know the optimal input of any technique, they may pay attention to their own use of that technique, but still be unable to use it. In order to investigate the strength of this mechanism, teachers will be asked to answer a questionnaire investigating the extent to which they wish they could use these different input factors when teaching. Teachers will be requested to answer these questions at endline.

We will use this variable to construct a measure of teachers' lack of training, which will be proxied by the difference between the extent to which teachers report they wish they could use a specific technique and the extent to actually use it (as observed by our surveyors). The greater the difference and the more problematic the lack of training is.

4.5 Sample size

Statistical power. We carry out statistical power calculations taking into account the fact that 17 schools and 68 language arts teachers enrolled in the study, representing an average of 4 teachers per school.

- Teacher-level outcomes: We carry out these power calculations under the assumptions that baseline covariates will allow us to explain 25% of the variation in our outcome variables. Our power analysis suggests that with a power of 80% we would be able to detect a minimum effect size of 0.50 to 0.60 standard deviations in teachers' behavior.[6]
- Student-level outcomes: We carry out these power calculations under the assumptions that baseline covariates will allow us to explain 28% of the variation in our outcome variables (estimated using Danish register data on past cohorts). Our power analysis suggests that with a power of 80% we would be able to detect a minimum effect size of 0.23 to 0.30 standard deviations in students' test scores. Parameter values on average class size, expected variance explained by each level's covariates, and inter-class correlations in test score data are estimated from Danish register data (past cohorts).[7]

[6] Calculations were performed in Optimal Design (Spybrook et al. 2011) choosing the individual-level randomization stratified by block design.

[7] Calculations were performed in Optimal Design (Spybrook et al. 2011) choosing the 3-level MS-CRT design with outcomes measured at level 1 (students' national test scores in reading), treatment at level 2 (teachers), and teachers nested in sites at level 3 (schools).

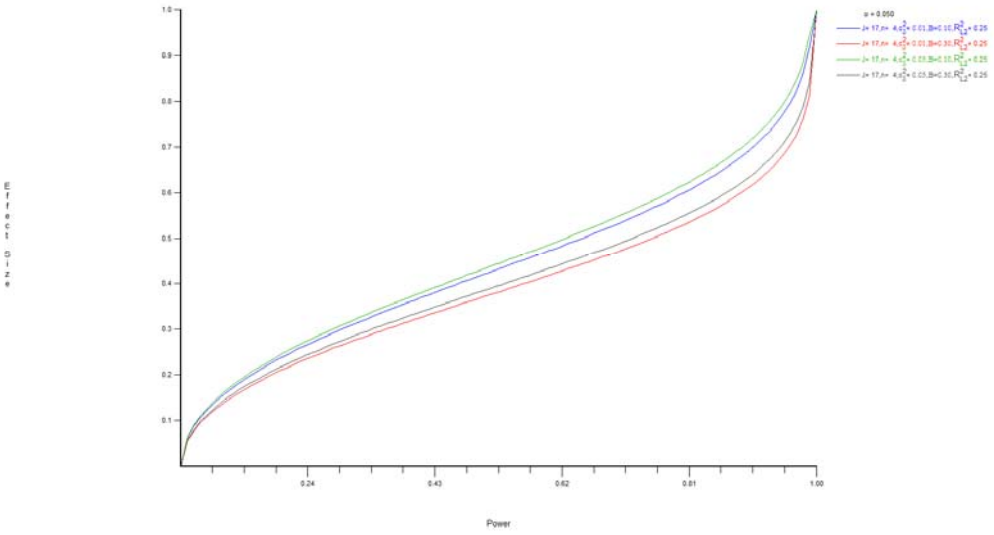

Figure 1 - Minimum detectable effect size as a function of power in a MS design (teacher-level outcomes)

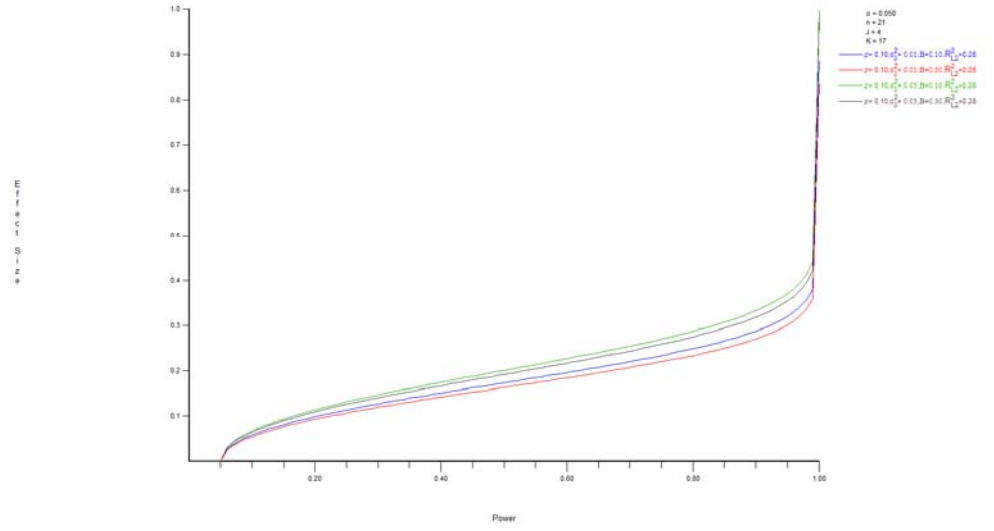

Figure 2 - Minimum detectable effect size as a function of power in a MS-CRT design (student-level outcomes)

References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034-1037.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics*, *128*(1), 1-51.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, *129*(3), 1311-1353.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Schwartzstein, J. (2014). Selective attention and learning. *Journal of the European Economic Association*, *12*(6), 1423-1452.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & TO, A. (2011). Optimal design plus empirical evidence: Documentation for the "Optimal Design" software. *William T. Grant Foundation.*

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485).