

**OBJECTIVES AND KEY RESULTS
IN THE CLASSROOM: A
FRAMEWORK FOR PRACTICE
DEVELOPMENT ON GOAL-
SETTING**

ANDRÉ LUÍS DA FONSECA

**A RESEARCH & DEVELOPMENT PROJECT
SUBMITTED FOR THE
MSc IN TEACHER EDUCATION 2018**

DEPOSIT AND CONSULTATION OF THESIS

One copy of your dissertation will be deposited in the Department of Education Library via the WebLearn site where it is intended to be available for consultation by all Library users. In order to facilitate this, the following form should be completed which will be inserted in the library copy your dissertation.

Note that some graphs/tables may be removed in order to comply with copyright restrictions.

Surname	Da Fonseca
First Name	André Luís
Faculty Board	Education
Title of Dissertation	Objectives and Key Results in the classroom: a framework for practice development on goal-setting

Declaration by the candidate as author of the dissertation

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I understand that the Department requires that I shall deposit a copy of my dissertation in the Department of Education Library via the WebLearn site where it shall be available for consultation, and that reproductions of it may be made for other Libraries so that it can be available to those who to consult it elsewhere. I understand that the Library, before allowing my dissertation to be consulted either in the original or in reproduced form, will require each person wishing to consult it to sign a declaration that he or she recognises that the copyright of this thesis belongs to me. I permit limited copying of my dissertation by individuals (no more than 5% or one chapter) for personal research use. No quotation from it and no information derived from it may be published without my prior written consent and I undertake to supply a current address to the Library so this consent can be sought.
3. I agree that my dissertation shall be available for consultation in accordance with paragraph 2 above.

Objectives and Key Results in the classroom: a framework for practice development on goal-setting

Abstract

Student motivation and engagement seem to be more important than ever considering information technology has rendered learning away from a classroom setting relatively commonplace. Also, recent research has provided evidence that the lack of non-cognitive skills may be harming academic results in developing countries such as Brazil. Motivated by these two circumstances, this study aimed to investigate the use of a modern goal-setting framework, Objectives and Key Results (OKR), as a tool to improve teachers' leadership skills and foster a more structured teacher/student dialogue about performance. A randomised controlled trial (RCT) was conducted to estimate the short-term effect caused on teachers by a lecture on OKR. Positive effects were detected, but the sample size prevents these results to be considered statistically significant. A single teacher OKR pilot was carried out to identify a manner of use for the framework in the classroom context. Pilot results look promising and probably justify the continuation of the research.

Table of Contents

Abstract.....	3
Introduction	6
Research aim.....	6
The quantitative phase	9
The qualitative phase.....	10
The importance of the research.....	11
Literature review.....	12
Do goals matter?.....	12
Objectives and key results in the classroom.....	18
Coaching for teacher leadership in Brazil	22
Methodology.....	28
Research questions	28
Question:.....	28
Sub-questions:.....	28
A mixed-methods approach.....	29
The experiment.....	30
The OKR pilot.....	43
Session part #1 - Framework.....	46
Session part #2 - Status.....	47
Session part #3 - Next steps	48
Scaling up	48
Ethics.....	49
Findings	52
A Randomised Controlled Trial	53
Participant flow	54
Recruitment and schedule	55
Baseline data.....	56
Outcomes and estimation.....	60
Availability of trial data	63
Limitations.....	64
Generalisability	66
Interpretation.....	66
A narrative case study of OKR in the classroom	69
Context.....	70
Phase #1 – Set	73

Phase #2 – Align	80
Phase #3 – Achieve.....	82
Limitations.....	85
Conclusion and Implications	87
The effect of a lecture on goal-setting.....	87
The practicalities of using OKR in the classroom	90
Non-cognitive skills and OKR.....	92
Implications and next steps	94
Bibliography	96
Appendix A	151
Appendix B	153
Appendix C	157

Introduction

Research aim

There is substantial evidence to suggest that present day ready availability of information is making possible a shift towards much more independence in the learning of students (see for example Katyal & Evers, 2004:367). Recent studies have indicated that the result of these more autonomous forms of learning has transformed the ways through which teachers are capable of causing a positive and relevant effect (Katyal, 2010:278). The importance of investigating the phenomenon described in the lines above, lies in the fact that teachers' ability to influence students within and beyond classrooms (Öqvist & Malmström, 2018:156), is said to play a "critical and central role in students' educational motivations" (Öqvist & Malmström, 2016:365).

The 2016 OECD report on low-performing students indicates that there is a growing number of students "trapped in a vicious circle of poor performance and demotivation" (OECD, 2016:3). Therefore, it seems to be important, in light of the more loosely coupled learning circumstances of the 21st century, that teachers are capable of maintaining their capacity of helping students to "achieve the behavioural and affective components of engagement" (Katyal & Evers, 2004:374).

According to the "self-determination theory" ("SDT"),

students have psychological needs to be "competent, autonomous, and related to others" (Deci & Ryan, 2012:1). Whenever these needs are not satisfied, motivation is said to be affected (Deci & Ryan, 2012:1). In a world where advanced information and communication technology ("ICT") corresponds to an excellent stimulus for students to self-determine their own learning (Lao, Cheng, Huang, Ku, & Chan, 2017:1169) and therefore fulfil their basic need for autonomy (SDT), if teachers are not aware that they may be playing a peripheral role as providers of knowledge (Katyal, 2010:282), they may end up limiting themselves to delivering the kind of instruction which blocks students' engagement other than fostering it (Katyal, 2010:282). At the same time, when teachers "support and encourage" students to "explore, initiate, endorse and engage in behaviours that are interesting and/or important for them", they end up facilitating the internalization of their influence upon students and fuelling students' motivation (Deci & Ryan, 2008:17).

Due to the above, I hypothesise that the "symbolic capital" or "professionalism" of teachers, in Bourdieusian terms (Schinkel & Noordegraaf, 2011:89), will start to be defined more in regards to their ability to cause a positive impact on students' non-cognitive skills, or, speaking freely, their ability to influence or "lead" the students, than in terms of the amount of first-order (e.g. Subject Content Knowledge (SCK) (Childs & McNicholl, 2007:1630) or even second-order knowledge

(e.g. Pedagogical Content Knowledge (PCK) (Shulman, 1986:9) that they may have accumulated during the course of their careers.

Based on the conjecture described above, the concept behind this dissertation can be summarised as follows: those more independent forms of learning, which were made possible by the ICT-driven ubiquitous access to knowledge, may be requiring teachers to enact a higher level of "leadership" (Gilley, Shelton, & Gilley, 2011:388; Öqvist & Malmström, 2016:371). The definition of leadership used here is the "capacity of teachers to influence and enable students to apply their abilities and resources to a given cause" (Öqvist & Malmström, 2016:367).

In practical classroom terms, this form of leadership contains many elements of the "student-centred teaching" or "student-centred learning" approach ("SCT") and vice-versa. Goal-setting, for example, is an element in the SCT approach (E. Lee & Hannafin, 2016:717) and is also considered a fundamental component in many accounts of leadership (Berson, Halevy, Shamir, & Erez, 2015:146; Fowler, 2018:188).

The research described herein leveraged this relationship to investigate the impact caused on teachers' level of self-efficacy for SCT by a teacher education intervention. The study encompassed two phases.

The first phase had a quantitative nature and focused on estimating the "Average Treatment Effect (ATE) (Deaton &

Cartwright, 2018:3)" caused by a lecture on goal-setting, delivered by the teacher educator, on teachers' self-efficacy for SCT. The second phase was a qualitative analysis of the teacher educator working with a teacher to implement a modern goal-setting framework within state-owned schools in Brazil.

The quantitative phase

The research's initial phase focused on evaluating a single lecture on goal-setting as a method to increase teachers' self-efficacy for student-centred teaching. As it frequently happens in many places, the format of teacher education interventions in Brazil is commonly limited to lectures and workshops (Marcondes, 1999:206). However, it is widely demonstrated in the literature that providing a "body of knowledge" to teachers typically fails to produce a lasting effect on behaviour (Kennedy, 2016:956; Korthagen, 2010:99). Although recognising the low efficacy of the format described above, this study considered this suboptimal reality a given fact and aimed to estimate the short-term impact of a viable intervention, in this case, a lecture, so to be able to usefully inform the addition of knowledge on goal-setting to existent teacher education programs. The quantitative approach adopted was the execution of a randomised controlled trial ("RCT") (Hacking, 1988:449) which involved a total of 32 (thirty-two) secondary teachers. This method was chosen due to

its prominence, justifiable or not, as the "gold standard" of credible evidence (Cartwright, 2011:1400). Also, besides providing quantitative information on the impact of a lecture on teachers' self-efficacy, the fact that a small RCT was used in such a way to evaluate a teacher education intervention may give birth to some insights on the applicability of randomisation in the field and may end up opening new avenues for research.

The qualitative phase

After the execution of the RCT mentioned above, a secondary teacher who participated in the experiment has volunteered herself to engage in a goal-setting pilot project. The pilot was about the co-creation of an educational version of one of the most successful goal-setting frameworks: the Objectives and Key Results ("OKR") (Wodtke, 2016:1). During this qualitative phase, I worked with this teacher for six weeks to apply OKR in her classrooms so to specify the ways through which OKR's "goal-setting conferences" (Gaa, 1973:42; Schunk, 1990:80) should be implemented in our context. We have outlined some of the characteristics of the OKR protocol when the framework is used in a classroom. We took into consideration the Brazilian K-12 national curriculum.

The importance of the research

As explained above, this dissertation investigates the use of goal-setting as an element of the knowledge base of teachers and teacher educators, accordingly to a conceptualisation of teachers as leaders (Öqvist & Malmström, 2016:365). Along these lines, the importance of the research described herein lies in two different aspects:

1. If verified to be viable and practical, the addition of knowledge on OKR to existent teacher education programs could be a powerful tool to increase, through the formation of teachers able to promote mastery goals (Ames, 1992:262), students' motivation to learn (Vedder-Weiss & Fortus, 2018:182). Also, recent quantitative research has suggested, drawing from data related to PISA exams, that Brazilian students may maintain a lower than expected level of motivation (Balart, Oosterveen, & Webbink, 2018:139). Such a result may be an indication that goal-setting frameworks have the potential to be highly beneficial when applied to the Brazilian educational context.
2. The use of a small-scale RCT to estimate the impact of an episodic and standalone intervention, such as a lecture, constitutes a novelty in the field of teacher

education and may set the stage for the insertion of such experiments into professional development ("PD") programs;

Literature review

Having given the rationale for this study and a brief outline of how the intervention was planned and carried out, this section begins by discussing the extent to which goal-setting matters in education. It then moves on to look specifically at the evidence for the particular goal-setting framework (OKR) that was introduced to teachers to develop their knowledge on goal-setting. Next, it is going to be discussed the possibilities of leadership coaching in the context of the Brazilian state-owned educational system.

Do goals matter?

Goals, or more specifically, achievement goals, refer, in the educational context, to students' reasons for undertaking tasks (Zong, Zhang, & Yao, 2017:342). The commonly accepted model of achievement goals establishes the existence of three goal types: mastery, performance-approach and performance-avoidance (Kaplan & Maehr, 2007:143). These performance-related types of goals are said to correspond to two different

motivational orientations. Performance-approach goals relate to students wanting to demonstrate high ability, and performance-avoidance goals refer to students trying to avoid the possibility of demonstrating low ability (Kaplan & Maehr, 2007:144). There has been some controversy regarding performance goals being detrimental within the educational context. Pintrich (2000) conducted a study involving one hundred and fifty 13-year old students (78 girls; 72 boys) and concluded that a focus on "approaching the competition and comparison" may not cause negative effects, in the case the student is also interested in mastering his or her schoolwork (p. 553). At the same time, Gonçalves, Niemivirta, & Lemos (2017) suggest that performance goals may cause "either neutral or negative motivational effects" in younger kids while causing "positive effects on college students" (p. 151). These same authors say that the goal orientations of students may vary, in different degrees, as they go thru schooling (ibid, p. 151). Zong et al. (2017), conducted a study involving 614 students from elementary schools in China and found that the prevalence of performance goals in that group of participants was positively correlated with the degree through which families were involved, inside the home, with student's education (p. 352). The same study showed that the occurrence of mastery goals becomes more frequent when parents actively participate in school activities (p. 353).

Different from what happens in the case of performance

goals, *mastery goals*, that is, goals "which focus on developing competence (e.g. "to learn")" (Marina S. Lemos, Gonçalves, Lens, & Rodrigues, 2014:1), are commonly considered to be beneficial. Ames & Archer (1988), for example, conducted a study involving 176 high-school students (85 girls and 91 boys) and verified that those students who emphasized mastery goals were the ones to demonstrate a "more positive attitude towards the class" and a preference towards "challenging tasks" (p. 260). More recent research has suggested that not all students who pursue the attainment of mastery use the same criteria (Tuominen-Soini, Salmela-Aro, & Niemivirta, 2008:252). Intrinsic criteria, such as, "the feeling of knowing and understanding", seems to be used by some students (mastery-intrinsic goals). Other students use "grades and feedback" for evaluating their level of mastery (mastery-extrinsic) (Tuominen-Soini et al., 2008:252-253). Marina S. Lemos et al. (2014) say that these mastery-extrinsic goals seem "to lay between mastery and performance goals" and propose the distinction of at least three performance goal subtypes: simple evaluation ("good grades"), ability, and normative ("competition") (p. 2). Those authors conducted a study which involved 134 middle school students and 423 high-schoolers, all coming from three state-owned schools located in the North of Portugal. According to them, data from this study indicated that performance goals of the subtype "simple evaluation", that is, goals where the student simply wants "to get good

grades" (ibid, p. 3) have a nature that is "very similar to that of mastery goals". Also, these goals "will probably not show the debilitating effects on students' motivation and achievement that have been associated with other types of performance-related goals" (ibid, p. 9). Such a result is of paramount importance in the context of this dissertation. It may represent an indication that the adoption of a goal-setting framework shall not require teachers to completely detach the sequence of objectives being elaborated in the context of the framework, from school's formal evaluation process. This possibility may make it viable a greater alignment with the expectations of the other stakeholders (school, families etc.).

From an educational perspective, the importance of goals also lies in the fact that goal-setting is a foundational process in the self-regulation of learning, which is defined by authors such as Zimmerman (2002) as a "self-directive process by which learners transform their mental abilities into academic skills" (p. 65). Usher & Pajares (2007) conducted a meta-analysis of data collected from 3,670 students in six different studies of motivation and found that students' self-efficacy for self-regulated learning are strongly and positively correlated with student's orientation towards the attainment of goals (p. 454). Schunk & Mullen (2013) studied the integration of mentoring with self-regulation and say that since "goal-setting triggers self-regulated learning" (p.

372), the existence of shared goals may ignite the "co-regulation" of learning within groups of individuals (p. 374). According to these authors, when a mentor, over the course of a year, works together with a student so to help him or her to achieve a sequence of goals (portfolio), teacher regulation goes down, and there is an increase in both co-regulation and student regulation (ibid, p. 374). It is significant the importance of such a result for the research described herein. It may properly characterise what happens to the self-regulation of students as they go thru the sequence of "teacher-student" goal-conferences (Schunk, 1990:80) which are inevitably promoted by the adoption of a goal-setting framework.

At the teacher's side, goals and goal-setting maintain close links with the student-centred learning/teaching approach (SDT). Elen, Clarebout, Léonard, & Lowyck (2007) say that the transition from a teacher-centred to a student-centred type of instruction includes goals being increasingly selected by the students, not by the teachers (p. 108). These same authors say that students associate the existence of "clear goals" with "high-quality education" (p. 109). Along the same lines, Förster & Souvignier (2014) state that teachers need "diagnostic information" on their students' learning progress to be able to provide "individualised instruction". These authors have conducted a study which involved 900 fourth-grade students from Germany and state that "engaging students in

goal-setting might be fruitful to foster motivation" (p. 99). However, they highlighted the fact that for goal-setting to be beneficial, younger students may need help to reflect upon success and failure (p. 99). Moeller, Theiler, & Wu (2012) conducted a longitudinal study on the relationship between goal-setting and achievement at the classroom level. From 2005 to 2009 they followed 21 teachers and their 1,273 students. The study took place in the State of Nebraska (USA) and included 19 state-owned and four local schools. In general terms, results indicated a "significant relationship between student's ability to set goals and language achievement in the Spanish language classroom" (p. 164). Also, authors say that student-centred teaching has positively correlated with higher achievement. Also, according to them the personalisation of the learning goals may also play an important role in the success of goal-setting in the classroom (p. 164). Indeed, during the research described herein, we have observed that the individualisation of instruction may be somehow induced by the implementation of a goal-setting method. In our case, the moment a teacher introduced the concept of each student defining and pursuing his or her goals, students started to demand the teacher to pay more attention to what they were doing individually, their ideas, doubts etc. While it is unclear at this point whether this observation can be generalised to other settings, it would be interesting to investigate the extent to which the adoption of goal-setting,

amongst other things, also brings about a new pedagogy.

Objectives and key results in the classroom

Locke & Latham (2002) state that while goal-setting theory focuses typically on motivation related to workplaces, its results are reliable across fields if the method is applied correctly (p. 714). Perhaps the most visible evidence to support the notion that goal-setting frameworks created within the organisational field can be, and frequently are, successfully used in education, is the case of the "Specific, Measurable, Attainable, Relevant, and Time-bound (SMART)" framework. This methodology was firstly proposed by Doran (1981) in a context that was purely business-related (p.35). Despite this origin of the SMART framework, Day & Tosey (2011) say that in the realm of secondary and 16-19 education in Wales and England, whenever the teachers exercise goal-setting, it is likely that they are using SMART (p. 515-517). The same authors characterise the adoption of a goal-setting framework as a way "to facilitate the interaction between student and tutor/teacher in an action-planning dialogue" (ibid, p. 516). Aghera et al. (2018) conducted a randomised trial to compare the use of SMART goals and its methods against the standard methodology, in the context of medical education. The RCT involved 80 medicine residents (p. 112). Data from the trial showed that the use of the framework had increased the number

of educational actions performed by participants (p. 117).

On the other hand, Day & Tosey (2011), analysed SMART accordingly to the criteria defined in self-regulation theory, and conclude that this framework's manner of use has a variable nature and this characteristic allows it to be applied in an instrumental way, that is, teachers don't necessarily "listen to learners" (p. 521) when applying SMART. Along the same lines, Sull & Sull (2018), say that SMART "ignores the importance of discussing goals throughout the year" (p. 2). The SMART framework defines what a goal should be but does not establish essential practical aspects such as, for example, who is responsible for determining everyone's goals, and the level of transparency which should be associated with each goal, amongst other elements.

The "Objectives and Key Results (OKR)" framework is defined by Doerr (2018) as being a "collaborative goal-setting protocol for companies, teams and individuals". The framework has a simple structure where *objectives* represent what the individual or group wants to achieve, and *key results* "benchmark and monitor" how *objectives* are going to be achieved (p. 12). Contrary to what happens in the case of SMART, OKR can be viewed as being a protocol. That means that the set of operational rules which regulate its implementation is also part of OKR (see for example Wodtke, 2016:5-7). In this sense, OKR can be understood as being a specific format to be used by practitioners to encode each goal (objectives + key results)

combined with the set of rules which define how these encoded goals are supposed to be created, assigned, monitored, processed, deleted etc. While it is beyond the scope of this dissertation to discuss such conceptualization in further detail, it is worthwhile outlining some of the rules which are part of the OKR protocol: (i) An *objective* is a single, qualitative and inspirational sentence (Wodtke, 2016:5); (ii) An individual is supposed to have created at least 50% of the OKRs which are assigned to himself; (iii) An *objective* should be stretch (Doerr, 2018:26), having a "50% chance of achievement at the time of setting" (Radonic, 2017:30); (iv) OKR attainment is not supposed to be directly linked to financial rewards (Zhou & He, 2018:321); (v) Key *results* are quantitative metrics or milestones (Sull & Sull, 2018:5); (vi) OKRs can and should be tracked, "revised and adapted as circumstances dictate" (Doerr, 2018:68); (vii) OKR achievement should be public and visible (Doerr, 2018:161);

When examining some of the OKR rules listed above, it becomes clear that much of what is defined resonates well with current accounts of self-regulation and self-discipline in learning (Zimmerman & Kitsantas, 2014:145). The rule which says, for example, that an *objective* needs to be expressed in the form of an "inspirational sentence", seems to be aligned with the notion that "learners need to possess or develop supportive, motivational beliefs" (Zimmerman & Kitsantas, 2014:145). As another example, the rule which establishes that

OKRs are supposed to be continuously tracked appears to be beneficial within the educational context, given that students are said to more easily sustain their intentions when they can monitor and control their performance (Zimmerman & Kitsantas, 2014:146). The recommendation that OKRs should not be associated with financial rewards may be correlated with the widely accepted notion that "mastery goals" correspond to the type of achievement goal which causes the most significant learning benefit for students (Kaplan & Maehr, 2007:143). In this sense, OKR achievement is not supposed to give any extra point or grade to students. Also, the rule which says that a *key result* needs to have a number (quantitative) is compatible with the observation that performance goals of the subtype simple evaluation, which some authors consider to be the "main goals pursued by students" (Marina S. Lemos et al., 2014:8), and mastery goals of the subtype extrinsic, both include the pursuing of "good grades". Therefore, in practical terms, this rule makes it possible the definition of a key result containing grades as one of its elements. Finally, the OKR rule which says that because "transparency seeds collaboration" (Doerr, 2018:49), OKR achievement needs to be publicly tracked, seems to be aligned with the notion that the existence of "opportunities to receive support from peers" is one of the factors which support the enactment of self-regulated learning (Mykkänen, Perry, & Järvelä, 2015:2).

Due to all the above, it looks promising the use of OKR

inside classrooms. The framework, or protocol, seems to address a significant number of aspects which are commonly associated with recent theories in education. Its successful implementation in a classroom may induce a higher level of "classroom leadership" (Greenier & Whitehead, 2016:79) amongst teachers.

Coaching for teacher leadership in Brazil

Komives et al. (2018) start their piece on the development of leadership by asking: "What kind of leadership is needed for a changing world?" (p. 23). Such question and all of its variants, have been occupying my mind since my Masters at Oxford was half completed, or more specifically since I had the opportunity to observe and analyse, in one of my assignments, a particular group of teachers from my hometown.

Those teachers worked for a program, called PEJA (Silva, Bonamino, & Ribeiro, 2012), which focuses on educating teenagers and adults who were left behind during compulsory education. What I have observed was a real tragedy. At that time, I was aiming to characterise and model the Subject Content Knowledge (Childs & McNicholl, 2007:1630) of teachers. In one of these occasions, I observed a Biology class in which the teacher was a Biologist and also a PhD candidate who was studying "insects" at the Federal University. After interviewing her, I thought that it would be safe to assume

that she knew a lot about her subject. Also, she was a seasoned teacher with more than 10 years of classroom experience; therefore, according to theories on the expertise of teachers (Berliner, 2004a: 207), I considered it very likely that she was in possession of at least some Pedagogical Content Knowledge (PCK) (Shulman, 1986:9).

One day, I watched this teacher making students copy content from the whiteboard to their notebooks during a four-hour lesson. She didn't ask anything nor demonstrated any interest in checking for student understanding.

That lesson observation was a moment of truth for me. Since then, I have been formulating the idea, in different forms, that maybe there is an under researched equivalence connecting accounts of leadership to the professionalism of teachers. In my view, that teacher was thinking in terms of "outputs" (tasks; products) not "outcomes" (results; benefits). There was no sense of responsibility related to the final result produced by the instructions she gave to the group of people that she was "leading" at that point. Her attitude corresponded, plain and simple, to a lack of classroom leadership (Greenier & Whitehead, 2016:79).

Öqvist & Malmström (2016) say that teachers who maintain a lower degree of classroom leadership end up not being sufficiently performance-oriented, participatory and supportive (p. 374). These authors have conducted a study which

involved interviewing 35 students. The study aimed to understand how teachers' leadership behaviours influence students' educational motivation. They conclude that "teachers' leadership activates and encourages students' educational motivation" (p. 376). Öqvist & Högström (2018) say that is part of the leadership of teachers the capacity to "distinguish and highlight technology in childrens' everyday lives" (p. 5). Souza & Batista (2018) conducted a study on the relationship between academic performance and the style of leadership of teachers. The study involved 119 students from elementary schools located in the State of Parana, South of Brazil. These authors indicated a higher prevalence of "teachers with an authoritarian leadership style" in the perception of the students with the lowest level of academic performance. Greenier & Whitehead (2016) say that teachers don't reflect much on the notion of "classroom leadership". These authors blame teacher educators not to discuss the topic with teachers "despite principles of leadership being strongly connected to the act of teaching itself" (p. 79). They also say that "leadership is manifest in every great teacher". The same authors conducted a study which involved 56 "native speaking English teachers", aiming to correlate these teachers' views on the characteristics of a competent language teacher, with same teachers' views about the principles of leadership (ibid, p. 84). The authors indicated that the small number of responses explicitly referencing the principles of

leadership was a function of teachers having a "behaviourist view of what teachers do when teaching" (ibid, p. 88). They end up saying that the knowledge base of teaching, as it is today, don't help teachers to "excite, inspire and build relationships with their students" (ibid, p. 88). Fowler (2018) says that the "competency to encouraging autonomy" is one of the abilities that a leader should have (p. 188). She states that a leader is supposed to "illuminate boundaries, then explore choices within those boundaries" (p.188). Katyal & Evers (2004) conducted a study in 3 schools located in Hong Kong. The study involved 14 teachers, 12 parents and 30 students (p. 369), and found teacher leadership to be "complex and multi-faceted". These authors say that teachers need to exercise a more "holistic form of "pedagogy" due to the new realities faced by schools (p. 381).

As it was briefly described above, the review of the literature seems to reveal the existence of a growing number of educational researchers who are starting to look at teachers' classroom behaviour thru the lens of leadership. Inspired by the notion that "teaching" and "leading" are in many ways, equivalent, I structured the OKR pilot that is part of this research, as a "leadership coaching" activity. My intention was not only to train a teacher on the specifics of the OKR method. I aimed to develop her identity as a leader as well.

A recent randomised study conducted by Bruns, Costa, & Cunha (2018) has revealed that "expert coaching" was capable of producing a "statistically significant 0.25 standard deviations increase in time on instruction" (p. 228). "Time on instruction" is a measure of the time spent by the teacher with the instructional tasks. In Brazil, this number is low for various reasons. No Brazilian teacher is said to average more than "66% of class time on instruction" (ibid, p. 215). This study took place in the State of Ceara, Northeast of Brazil, and serves as an indication that coaching initiatives are well accepted within the local educational system. Kraft & Blazar (2017) say that the US is seeing an increase in the number of districts which concluded that "teacher coaching" constitutes a better alternative in comparison to lectures and short-term workshops (p. 1034). These same authors say that the majority of coach programs are based on an "individualised feedback process" (p. 1036). They also say that the typical structure of a coaching process "consists of classroom observations followed by targeted feedback" (p. 1036).

In the context of the OKR pilot, the proposed structure included classroom observations (video) but was centred on the analysis of the teacher's discourse. Indeed, Yarborough (2018) highlights the similarities between coaching and the psychoterapeutic process (p. 50). He adds that different from psychoterapy, "coaching tends to focus more on present and

future states". In the practical case, I was interested in understanding the reality of the OKR implementation. However, I was equally interested in analysing teacher's interpretation of such reality. Because of this, I ended up observing just 20% of the lessons but discussing and examining together with her, 100% of her discourse on what was happening in the OKR pilot.

Although not explicitly, I based the last part of my interaction with the teacher on the first two stages of the Leadership Identity Development (LID) model (Komives et al., 2018:24). The first stage of the LID model is "Awareness" (p. 25). I found this stage of paramount importance since teachers typically don't see themselves in a leadership position. So, the first thing I did was to name the teacher who I was working with, the "OKR leader". We discussed different leadership styles during this stage. I asked her about the inspiring leaders she had the opportunity to meet in her life. We discussed different topics related to leadership but always maintained our focus on what was happening in the OKR pilot. The second stage of LID is "Exploration/Engagement" (ibid, p.25). During this stage, I asked her about the skills and characteristics that she already had that she considered beneficial or detrimental for the success of the OKR pilot. We discussed, for example, her tendency to always tell people what to do. We debated the compatibility of that kind of action

with the need to induce students to define their own goals and did some role play on Skype to address that specific issue.

In general terms, the coaching method seems to be appropriate to help teachers implementing OKR and also to improve their leadership skills. In the case the OKR pilot grows in size, it may be necessary the creation of more structured, perhaps web-based, coaching program (Kraft & Blazar, 2017: 1034).

Methodology

Research questions

Question:

In what ways and to what extent is a teacher education intervention on goal-setting an effective tool to shape instruction according to the principles of current educational reform?

Sub-questions:

What is the effect of a teacher education intervention on goal-setting on teacher's self-efficacy for student-centred teaching? (Quantitative phase)

What do teachers perceive as being the key issues and challenges for teachers with no experience with goal-setting to learn and efficiently apply the OKR framework? (Qualitative phase)

A mixed-methods approach

Johnson & Onwuegbuzie (2004) call pragmatism the “philosophical partner of mixed methods research” (p. 16). My inclination towards a “practical and outcome-oriented method of inquiry that is based on action” (p. 17), perhaps derived from my previous identity as a software programmer, explains my decision to adopt a mixed methods approach in this research. I suspect that a “class of research” which is, in the words of those authors, “expansive and creative”, “not limiting”; “inclusive”; “pluralistic”; and “complementary” (p. 17), has a better chance to make a real contribution to the development of education, particularly in developing countries. Along these lines, it should be noticed that although the research described herein can be said to be predominantly quantitative, given that the execution of an RCT has a major role in it, the randomised experiment itself can be viewed as a “combination of numerous conceptual steps and research procedures” (Mark 2015:17) some of which having a qualitative nature. Also, as it was explained in previous sections, this mixed methods research can be formally subdivided into two well-defined

phases: (i) The experiment; (ii) The OKR pilot;

The experiment

As explained above, the first phase of this research can be said to be predominantly quantitative. However, Mark (2015) states that the definition of the research question is a relevant element of what he calls "critical multiplism" (p. 18). Indeed, the quantitative research question which I aimed to answer in this research, that is, "What is the effect of a teacher education intervention on goal-setting on teacher's self-efficacy for student-centred teaching?", was inspired by "qualitative observations" (p. 18) which preceded the idea of an RCT. Before even thinking about the use of an experiment to estimate the impact of an teacher education intervention, I had previously noticed, through a set of semi-structured interviews which related to previous assignments, that secondary teachers in my context where not sufficiently "focusing on their students" (Kilday, Lenser, & Miller, 2016:62) despite their espoused beliefs. Because of this observation and drawing from literature which connects teachers' self-efficacy, that is, their "ability beliefs for completing a given task" with "classroom practice" (Kilday et al., 2016:62), I asked myself whether a teacher education intervention on one of the elements commonly associated with

the "student-centred dialogue with tutors" (Bullock & Wikeley, 1999:19), that is, goal-setting, would increase, at least momentarily, their self-efficacy for more student-centeredness in their teaching. The idea of using an RCT sounded like an exciting and novel method to try to answer this interrogation.

RCTs are designed for assessing the extent to which a "potential causal variable makes a difference in a possible effect variable" (Mark, 2015:15). The former is frequently called the "independent variable", and the latter is regarded as the "dependent variable" (Mark, 2015:15). In this research, the independent variable was defined as attending a lecture on goal-setting. In this case, provided that I was the one responsible for giving the lecture and considering that I was in full control of its content, it seemed safe to assume that my research procedures were in "fact manipulating the independent variable construct of interest" and not other constructs (Mark, 2015:19). In other words, I intended to run a randomised experiment to measure the impact of a lecture on goal-setting and, from the perspective of the independent variable being what it was supposed to be, this is what I did. Therefore, there is not much to say about the lecture itself other than stating that it was about the OKR framework (Doerr, 2018) (Wodtke, 2016). The lecture was about OKR being used in a classroom by a computer science teacher (see Appendix A for a sample of the slides used in the lecture). Regarding the dependent variable, that is, the outcome which I was trying to

measure, the methodological approach adopted in the research requires more detailed consideration. The lecture took place in a teacher education centre run by the Estate of Rio de Janeiro. The Secretariat of Education invited a total number of 32 (thirty-two) secondary teachers. 96.9% of the teachers attended the invitation (n = 31). They were told that an educational researcher would be given a lecture on goal-setting as part of his (mine) research and that their participation was 100% optional and would require each of them to provide explicit consent and fill a set of questionnaires both before and after the lecture. Teachers were equally invited with no specific information supplied to any of them. However, for the RCT to take place, I had to randomly assign each teacher to one of two equally sized groups, Treatment and Control. Each of these groups ended up attending the lecture in a different moment during the same day.

Heckman (1991) says that for randomisation to create two similar groups which differ only in the fact that one of them had access to the treatment, neither "potential applicants" nor "program administrators" should be given information about the "possibility of randomisation" (p. 16). The requirement for participants not knowing that they were selected for the Treatment group is simple to understand: there is a risk that the methods used to measure the dependent variable end up measuring a combination of this variable plus any psychological

effect caused by knowing that the individual was randomly selected to participate in "an experiment" (Scriven, 2008:13). The way I dealt with this constraint was by divulging the lecture as being a normal teacher education intervention in which questionnaires would be used to measure teachers' content assimilation and other related aspects. Indeed, the intervention was structurally equivalent to the kind of training event those same teachers are used to attend. Finally, the second requirement presented by Heckman (1991), which says that program administrators should not know that applicants are randomly selected to "receive treatment", does not apply in the case of the research described herein since all teachers end up attending the lecture. Therefore, there was not an incentive for program administrators, in this case just myself, to screen more persons from a specific group so to have more of these individuals selected for the Treatment group by the randomisation process (p. 17).

RCT's dependent variable was defined as self-efficacy for student-centred teaching. To measure this construct, I used a modified version of the "Self-Efficacy for Student-Oriented Teaching (SE-SOT)" 12-item questionnaire shown in (Kilday et al., 2016: 68). SE-SOT items are presented with "a single item stem 'I am _% certain that I can..' followed by the instructional task" (Kilday et al., 2016:66). Table 1 bellow presents the 11-item modified version of SE-SOT (SE-SOT*) used

in this research. The only modification from the original SE-SOT is the exclusion of one of the two somewhat similar questions related to helping students recognising the value of the material.

Table 1 - SE-SOT* questionnaire	
1	I am _ % certain that I can present content that students relate to other subjects
2	I am _ % certain that I can provide a rationale to make academic tasks relevant
3	I am _ % certain that I can recognise students' thoughts in my explanation of learning tasks
4	I am _ % certain that I can help students to recognise how the material has value for learning
5	I am _ % certain that I can create authentic learning opportunities to make knowledge relevant
6	I am _ % certain that I can help students to monitor their performance appropriately
7	I am _ % certain that I can help students to identify strategies to support their learning
8	I am _ % certain that I can model strategies that students can use in other classes beyond the academic year
9	I am _ % certain that I can assist struggling students in identifying strategies to support their learning

10	I am _ % certain that I can provide opportunities for students to set their own learning goals
11	I am _ % certain that I can create opportunities for students to make choices about their learning

The essence of an RCT is to be able to compare a “baseline measure” (White, 2013:31) of the *dependent variable* with a post-treatment, or, in this case, post-lecture, measurement of the same variable. Also, it is a requirement of such an experiment the submission of the *Control* group, that is, those who did not receive treatment, to the same post-treatment measurement process. Such a task is necessary to produce a reliable evaluation of the counter-factual, i.e., an estimate of what would have happened should the *Treatment* group was not exposed to treatment (White, 2013:31). However, due to this research’s short term, some of these measurements were supposed to be separated by just a few minutes. As an example, I wanted participants from the *Control* group to fill the post-lecture questionnaire as soon as they left lecture’s room at the very beginning of the event, so to avoid “spillovers” (White, 2013:37). To do that, I had to deal with the fact that I just possessed a single questionnaire, SE-SOT*, to measure the *dependent variable*. Therefore, given that it does not make sense to ask a teacher to fill the same questionnaire twice in such a short time span, I had to find a workaround for the problem of constructing two similar but different measuring

instruments. The way I went about this issue was by creating an alternative item stem. In addition to the original "I am _% certain that I can..", I created similar questions with the alternative stem "When it comes to being capable of.., I am a _ (1-10) teacher". Also, to prevent differences in teachers' interpretation of the two question formats to affect results, I created code (R language) to randomly select, for each item of each questionnaire, one of the two forms. By doing this, I managed to create two questionnaires containing questions in both formats, and teachers ended up participating in two equivalent but "apparently" different surveys. Tables 2 and 3, bellow, show the final version of each questionnaire used in the research. It should be noticed that I did not pilot the two questionnaires before using them in the RCT. However, since each group, Treatment and Control, ended up answering the two questionnaires, and, given the fact that the estimation of the average treatment effect includes a "difference of differences", it was not necessary to guarantee the equivalence of the two instruments. Whatever the difference between the two, this difference equally affected both groups and, in this sense, it was cancelled out during the estimation process (See Appendix B for a sample of the real questionnaires used in the RCT).

1	I am _ % certain that I can help students to identify strategies to support their learning
2	I am _ % certain that I can model strategies that students can use in other classes beyond the academic year
3	I am _ % certain that I can provide opportunities for students to set their own learning goals
4	I am _ % certain that I can help students to recognise how the material has value for learning
5	I am _ % certain that I can provide a rationale to make academic tasks relevant
6	I am _ % certain that I can create opportunities for students to make choices about their learning
7	I am _ % certain that I can help students to monitor their performance appropriately
8	I am _ % certain that I can assist struggling students in identifying strategies to support their learning
9	When it comes to being capable of presenting content that students relate to other subjects, I am a _ (1-10) teacher
10	When it comes to being capable of creating authentic learning opportunities to make knowledge relevant, I am a _ (1-10) teacher
11	When it comes to being capable of recognising students' thoughts in my explanation of learning tasks, I am a _ (1-10) teacher

1	I am _ % certain that I can recognise students' thoughts in my explanation of learning tasks
2	When it comes to being capable of modelling strategies that students can use in other classes beyond the academic year, I am a _ (1-10) teacher

3	When it comes to being capable of helping students to identify strategies to support their learning, I am a _ (1-10) teacher
4	When it comes to being capable of helping students to monitor their performance appropriately, I am a _ (1-10) teacher
5	When it comes to being capable of helping students to recognise how the material has value for learning, I am a _ (1-10) teacher
6	When it comes to being capable of assisting struggling students in identifying strategies to support their learning, I am a _ (1-10) teacher
7	When it comes to being capable of creating opportunities for students to make choices about their learning, I am a _ (1-10) teacher
8	When it comes to being capable of providing opportunities for students to set their own learning goals, I am a _ (1-10) teacher
9	I am _ % certain that I can present content that students relate to other subjects
10	I am _ % certain that I can create authentic learning opportunities to make knowledge relevant
11	When it comes to being capable of providing a rationale to make academic tasks relevant, I am a _ (1-10) teacher

As it was mentioned above, 32 secondary teachers were invited, and 31 attended the event on February 22nd, 2018. The session started with me explaining that I was there to research goal-setting as part of my Master's degree at Oxford. I then went on to tell that their participation was 100% optional and that if anyone decided to leave the room at any moment, he or she could do that without giving any explanation. I told them that they would be asked to fill two 11-item questionnaires and that the confidentiality of the collected data was going to be protected by Oxford University's codes of conduct. Then

my assistant gathered a signed *Letter of Consent* from each teacher and assigned a number from 1 to 32 to each of them. I then asked the whole group of teachers to fill the pre-lecture questionnaire shown in Table 2. After everyone had filled the survey, I told the remaining 30 participants, for the first time, that due to the amount of groupwork I intended to promote during the lecture, I had to subdivide the group into two halves and should initially give the talk to only 50% of the participants. At this point my assistant used a list of 15 integers, which were previously randomly selected from the interval [1-30] (R code) and invited each teacher who was assigned to one of those numbers to follow her to the room next door so to have coffee and biscuits while waiting for their turn to attend the lecture. While my assistant was outside having coffee and chatting with those teachers, she took the opportunity to ask that group, from now on known as the *Control* group, to fill the post-lecture questionnaire (counterfactual).

Meanwhile, I was in the other room giving the lecture to the first group of teachers, the *Treatment* group. Fifty minutes later I ended the talk and groups switched places. Those leaving the main room, the *Treatment* group, were then also invited to have coffee and biscuits and were asked by my assistant to fill the post-lecture questionnaire before leaving premises. I then moved on to giving the same lecture

on goal-setting to the *Control* group so to make sure everyone had the opportunity to benefit from the same instruction.

The sequence of steps related to the definition and operationalisation of experiment's dependent and independent variables, as described above, is represented in Figure 1 below.

After having conducted the operational steps required to measure the dependent variable before and after the intervention, it was then necessary to analyse data. In general terms, it could be a conclusion of such analysis that the experiment ended up not revealing a statistically significant result. Indeed, one of the most fundamental questions related to the execution of an RCT is "How large a sample do I need" so to be able to "detect" a meaningful effect (White, 2013:38). An experiment's sample size needs to be sufficiently big to make differences occurring "by chance" to be rare (Stokes, 2014:180). At the same time, the number of participants should not be "larger than necessary" so as not to waste resources (Stokes, 2014:180). In this respect, it needs to be noticed that the experiment described herein was designed with the intention to have it inserted into an existent teacher education program to produce evidence on the effectiveness of an intervention which is already part or will be made part of such a program.

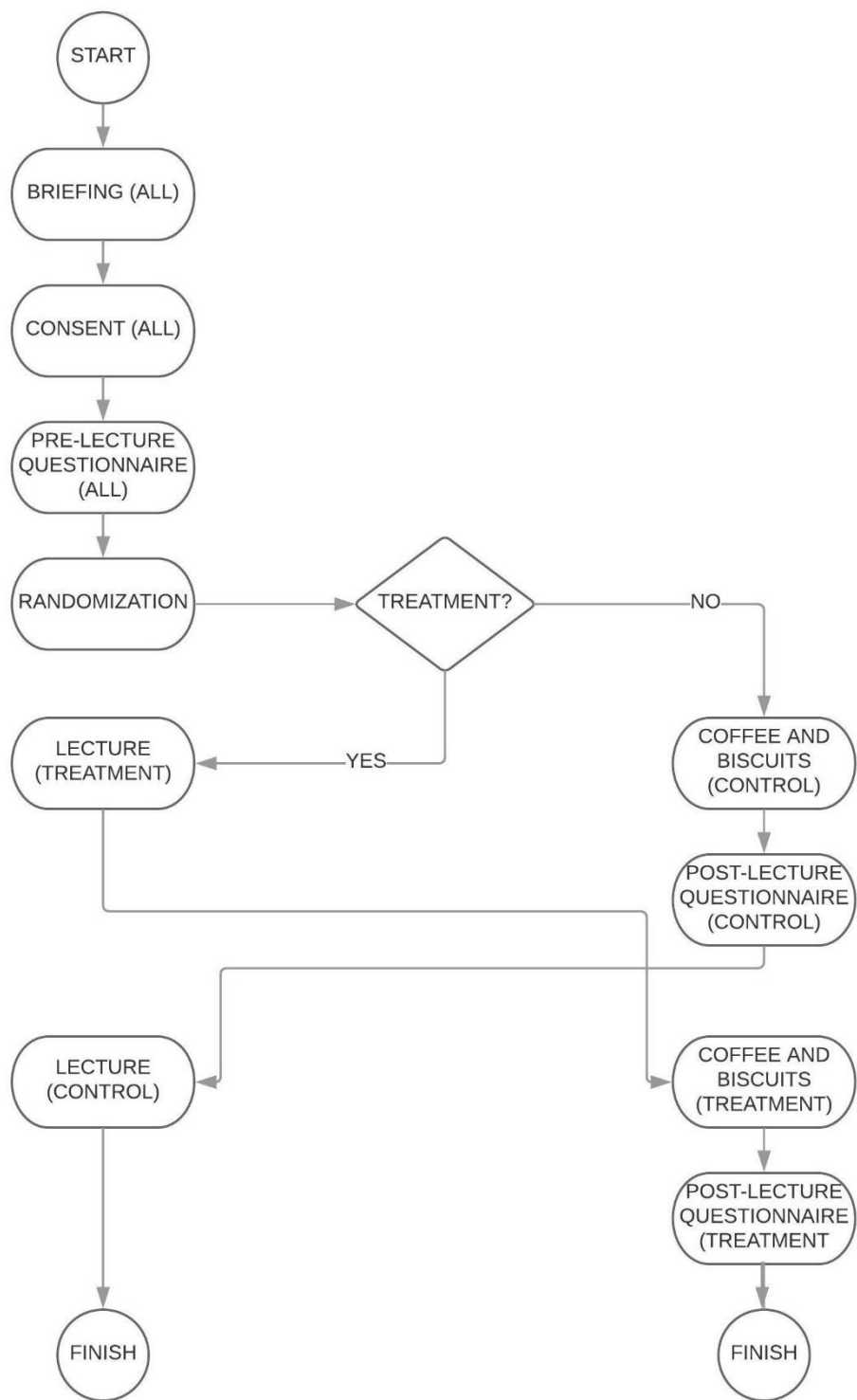


Figure 1 - RCT's flow diagram

Therefore, in such conditions, the size of the sample ends up being limited to the size of the program's cohort. Should this size ends up being smaller than necessary to make it statistically significant the observed differences (Stokes, 2014:180), that is, in the case the "power" (Cook & Hatala, 2015:74) of the experiment ends up being small, the same experiment needs to be repeated with different cohorts. Conveniently, meta-analysis can be used to combine results from a sequence of "low powered" experiments, to increase power (Duflo, Glennerster, & Kremer, 2007:3928). At the same time, the characteristics of the experiment described herein do allow for the "taking of repeated measures of outcomes" to "obtain more precise estimates of effects" (McKenzie, 2012:2010). Indeed, this study was intentionally designed to measure the effect of an intervention which was going to be offered to teachers as part of an ongoing teacher education program, regardless of the existence of the RCT. Because of this, it seems safe to assume that an RCT of this kind could be "embedded" into the program and repeated in multiple years so to, at a little cost, circumnavigate the sample size issue.

Due to the explained above, this research's data analysis treated the sample size, and therefore, the statistical power of the experiment, as given data. However, it should be noticed that showing the "observed power" that is, reporting a post-hoc power analysis based on actual sample size (30) and

experiment's outcome, is useless due to statistical reasons which are well beyond the scope of this dissertation (Hoenig & Heisey, 2001:20). Instead of doing this, I limited myself to showing in the findings how probable it was for the experiment to not detect a large enough effect in the case there was one, provided that the sample size was what it had to be. Also for illustrative purposes, it will be presented what would be the recommended sample size, should we wanted to set the experiment's minimum power at 80% (Cook & Hatala, 2015:74). The metric used to present results was the "Standardized Mean Difference (SMD)" because it facilitates the comparison of results across studies (Cook & Hatala, 2015:74). Lastly, it must be said that the guidelines described by Schulz, Altman, Moher, & Group (2010) were followed to report results (p. 727) whenever applicable.

The OKR pilot

The second and qualitative phase of this research relates to the study of a real attempt to implement OKR in the classroom. It aimed to understand the extent to which it would be possible for a secondary teacher, with no previous experience with the topic, to ignite and lead a new goal-setting process via the use of the OKR framework. Also, it intended to observe whether the implementation of OKR would induce teachers to assume the role of a leader, as defined in

the Introduction. The methodology employed in this qualitative phase almost required my insertion into the subject of study. Indeed, the novelty associated with the use of OKR in the classroom, that is, the fact that there was no previous experience from which I could draw practical information, obliged me to collaborate with the teacher who was piloting the framework closely. In this sense, it can be said that there was a lot of "co-creation" taking place during this phase. Therefore, it seemed virtually impossible to differentiate in this teacher's discourse, what was "her" voice from what was "our" voice. Having said that, I am cognizant of the risk that qualitative research, especially when it uses none of the devices which are commonly adopted by quantitative researchers to control bias, ends up becoming what Thirsk & Clark (2017) call "just well-composed rhetoric of a well-meaning, but biased, researcher's opinion" (p. 4). In this regard, it seems prudent to recommend that the qualitative method utilised in this research should be understood as being just a "pragmatic constructivist" case study (Harrison, Birks, Franklin, & Mills, 2017:10). As such, it aimed to produce nothing more than a "rich, holistic description that illuminates one's understanding" (Harrison et al., 2017:10) of an OKR implementation within a Brazilian secondary classroom.

Regarding the risk of bias in qualitative research, Galdas (2017) says that whoever carries out this type of study is "an

integral part of the process and the final product, and separation from this is neither possible nor desirable" (p. 2). In this sense, other than claiming that this qualitative work "reflects objective, opinion-free neutrality" (ibid, p.2), I adopted a critically reflective stance on my own preconceptions and focus (ibid, p.2) and maintained "methodological coherence", as defined in (Morse, Barrett, Mayan, Olson, & Spiers, 2002:18).

The research question which is associated to this qualitative phase is "What do teachers perceive as being the key issues and challenges for teachers with no experience with goal-setting to learn and efficiently apply the OKR framework?". To be able to offer an answer to this question, I asked the Secretariat of Education to send an email to the teachers who have attended the lecture on goal-setting, inquiring whether any of them would be interested in working with me to pilot the implementation of OKR in her or his classroom. Almost immediately, a 48-year old English as a Foreign Language (EFL) teacher, *Karen* (her real name was omitted to maintain confidentiality), replied to the email saying that she not only wanted to participate but was already trying to apply some of the concepts which I talked about during the lecture. The first of my encounters, from this point on called "sessions", with *Karen*, happened at the same teacher education centre where the OKR lecture took place. Further

sessions were conducted mainly via Skype to reduce costs. Each session had the following three-part structure: (i) Framework; (ii) Status; (iii) Next steps;

Session part #1 - Framework

A typical session used to start with me presenting an OKR-related concept, such as, for example, "OKRs are used to focus a group or individual on a bold goal" (Wodtke, 2015:4). In this case, after I finished presenting the concept, the session then assumed the characteristics of an open discussion in which *Karen* and I debated the notion of "a bold goal" applied to the task of inducing students to define their own goals about a certain topic of the curriculum. Many questions used to arise. Is being "bold" equivalent to being "difficult"? Is the notion of a "bold goal" supposed to vary as the student progresses through the year? How is a teacher (leader) supposed to communicate the fact that she expects an increasing level of "boldness" as students go from goal conference to goal conference? As it can be inferred from this list of questions, *Karen* and I have collaborated to adapt OKR, a framework which has its roots in the organisational field, to the classroom context. By acting this way, I ended up finding myself deeply involved with the process which I was aiming to study. However, the opportunity to co-create a new version of the OKR framework has given me the chance to witness a form of teacher growth

which appears to be under researched in teacher education: teachers' acquisition of instructional leadership skills (Öqvist & Malmström, 2018:159). Excerpts from these rather rich discussions were extracted. I consider them to characterise some of *Karen's* main points of concern regarding the adaptation of OKR to an educational setting. A sample of these fragments is presented and discussed in the Findings.

Session part #2 - Status

After debating the OKR "topic of the week" as described above, *Karen* and I used to move on to examining the status of the OKR implementation in her classroom. From a methodological standpoint, this gave me the opportunity to identify some of the hurdles which the "average teacher" would probably face when engaging with a real project. We discussed ways to circumnavigate every obstacle, which means that more than often the evolution of her practical work has counted on my input as a supervisor. It was possible to identify a small set of areas which need to be prioritised in further initiatives. These areas corresponded to roughly 80% of the issues and will be described in the findings.

Session part #3 - Next steps

I decided not to use OKR to supervise the teacher who was working with me in this study. One of the most prominent features of the original framework is its ability to “cascade” (Doerr, 2018:13), which means that, for example, managers’ OKRs may depend on the OKRs of employees. This feature signifies that OKR has the potential to be used as a unified language to talk about goals within an organisation or school. Due to this characteristic, I could have asked *Karen* to assign OKRs to herself so to better structure her attempt to adopt the framework. However, the novelty of the subject and a perception that perhaps the interaction “supervisor-teacher” would be different from the interaction “teacher-student” which we were struggling to characterise, made me decide not to explore the possibility to “model” in myself the behaviour which I was trying to induce *Karen* to assume. In this regard, without necessarily adopting any framework to define future goals, sessions used to end by the teacher and I merely planning for the next steps.

Scaling up

Karen was not the only teacher who worked with me in this study. Motivated by the results that she managed to achieve with her initial implementation, she took the initiative to

contact one of her colleagues, *Gisele* (real name omitted to maintain confidentiality), a 40-year old secondary science teacher. With an unexpected level of enthusiasm, *Karen* "sold" the OKR concept to *Gisele* by emphasising that, in her view, the OKR framework facilitates assessment. Indeed, assessment facilitation was a significant thing in the opinion of *Karen*. A comprehensive description of this aspect will be done in the Findings. The methodology used with *Gisele* was the same used with *Karen*, as described above. Due to time constraints, the description of my work with *Gisele* was left for future work.

Ethics

Because of the ethical issues that may arise as a result of this research, a CUREC form explaining the ethical measures which I have taken was completed. Data collection has started after I have received ethical approval. The 30 participants of the RCT described herein have granted written consent from an informed position. The teacher who was involved with the OKR pilot was one of these RCT's participants. I did not ask her to sign a second consent form before initiating the pilot. She was told that the totality of my interaction with her during the pilot was being analysed as part of the work related to this dissertation. To this day, I have not received any complaint from her nor any of the other RCT participants. Due to confidentiality reasons and file size, just a small sample

of the signed consent forms are shown in Appendix C. Originals will be safely preserved for future access.

RCT participants were fully briefed before the beginning of the experiment and were told that their participation was 100% optional. In fact, after the briefing one of the teachers decided that she did not want to participate and left the room without having to disclose any motive.

I recognise that the ethical aspects of RCTs in the social sciences are controversial in many forms. However, most of the criticism goes in the direction of saying that it is unethical to offer to just a portion of a population, treatment that is known to be or supposed to be beneficial (e.g. "The experiment aims to provide credible evidence for the effectiveness of the parachute. Half of you will be randomly selected to jump without one for us to observe the counterfactual."). In the case of the RCT described herein, both groups attended the same lecture on the same day. Therefore, the ethical issue described above does not seem to apply in this case.

However, a second ethical aspect which is undoubtedly more subtle comes from the need or desire to make the RCT entirely blind (Scriven, 2008:13). Ideally, participants should not know whether they are receiving treatment while other people are not. The problem is that in experiments of this type, participants end up knowing this "because they have been told it in order to meet local interpretations of

constraints on the use of human subjects" (Scriven, 2008:13). The way I dealt with this problem leveraged on the symmetrical aspect of the experiment, that is, the fact that everybody ended up receiving treatment on the same day. Because of this characteristic, I decided to not disclose, before the end of the experiment, information on who was part of Control and who was part of the Treatment group. I decided to do so under the straightforward interpretation of what says the "Ethical Guidelines for Educational Research" document which is published by the British Educational Research Association. Excerpt bellow shows what this document says in regards to the "Openness and Disclosure" principle (British Educational Research Association, 2011:6):

*The securing of participants' voluntary informed consent, before research gets underway, is considered the norm for the conduct of research. Researchers must, therefore, avoid deception or subterfuge unless their **research design specifically requires** it to ensure that the appropriate data is collected or that the welfare of the researchers is not put in jeopardy.*

Given that the omitted information caused absolutely no impact on any participant and provided my research design (RCT) required this minor degree of blindness to ensure that the appropriate data was collected, I claim that my decision did not harm my full compliance with the BERA guidelines (British Educational Research Association, 2011).

Findings

Wisdom, Cavaleri, Onwuegbuzie, & Green (2012) say that mixed methods may be an adequate approach when data generated by either quantitative or qualitative analysis does not sufficiently illuminate the topic so to allow for a proper understanding (p. 722). In the case of the research described herein, the primary objective was to investigate the use of a goal-setting framework as a tool to change teacher behaviour in such a way that teachers become more capable of positively affecting student's motivation and engagement. As we have seen in previous sections, from a teacher education perspective, such a goal could be divided into two separate objectives. Firstly, it was considered essential to know whether a typical intervention on the topic was capable of winning teachers' "hearts and minds" in such a way that we were able to measure an impact in their set of beliefs. Secondly, I considered it crucial for the acceptance of the method as a viable alternative, the assessment of the effort which corresponds to teachers trying to use the framework for the first time. The combination of these two different objectives made a clear indication towards the use of a mixed-methods approach. Detecting a "difference" in beliefs was properly addressed by an RCT, a quantitative method. At the same time, the description of the difficulties associated with implementing

OKR was structured as a "narrative case study", a well-known qualitative research method (Brandell & Varkas, 2001:294). Also, it should be mentioned that authors such as Wisdom et al. (2012), consider it of paramount importance that "reports from mixed methods studies transparently discuss the methodological components integral to the conduct of studies" (p. 722). Because of this, the lines below will present the findings from each phase of this research, quantitative and qualitative, in accordance to the guidelines which are commonly regarded as "standard" in each of the two situations. Although the two phases involved the same population of participants and corresponded to the same study, I decided to leave, for the sake of rigorousness, any connection between the two phases to be discussed in the Conclusions.

A Randomised Controlled Trial

Researchers from the biomedical field have developed a set of guidelines related to the reporting of RCTs in their research area (Grant et al., 2013:596). These guidelines, published by the CONSORT group (www.consort-statement.org), have not been fully adopted by many social researchers, amongst other reasons because the recommendations "may not be wholly adequate for social and psychological interventions in their current form" (Grant et al., 2013:596). Luckily, "a

collaboration of the Centre for Evidence Based Intervention at the University of Oxford, the Centre for Outcomes Research and Effectiveness at the University College London, and the Institute of Child Care Research at Queen's University Belfast, in association with the CONSORT group", has developed a CONSORT extension for social and psychological interventions, called CONSORT-SPI (Grant et al., 2018:2). This extension is supposed to become the "de facto" standard in reporting trials within the social sciences arena.

Motivated by the above, the lines bellow present the results of the RCT contained in this research in a way that is entirely complaint with the CONSORT-SPI 2018 guidelines (Grant et al., 2018:11-13).

Participant flow

A heterogeneous group of seasoned secondary teachers (n = 32) was invited via email by the Secretariat of Education of the Estate of Rio de Janeiro to attend a teacher education session on "goal-setting" which was scheduled to take place on February, 22nd, 2018. The session happened within a state-owned teacher education centre. All teachers but one attended the event (non-enrolment rate = 3.13%). Once within premises, these teachers, now called "participants", were briefed regarding the objectives of the session, ethical aspects, and

regarding the fact that the group was about to be subdivided into two halves so to facilitate group-work. Participants were told that all of them were guaranteed to attend the lecture during that same day and were asked to give explicit consent. Thirty participants (attrition rate = 3.22%) ended up giving consent. Those participants were then randomised to create two sub-groups: Treatment (n = 15; 50.0%) and Control (n = 15; 50.0%)

Recruitment and schedule

Recruitment happened in February 2018 (email invitation). All other activities of the trial took place on a single day (February 22nd, 2018). On that day, after the initial briefing, consent sought and randomisation stages (20 minutes), participants from the two sub-groups were asked to fill up a questionnaire so to provide baseline data (10 minutes). The "intervention" was a lecture on goal-setting (50 minutes). While the Treatment group was receiving the intervention, the Control group was outside the room filling up the endline questionnaire and waiting for its turn to receive the treatment (see the Methodology section). At the end of the first lecture, sub-groups switched places and then the Treatment group was invited to fill the endline questionnaire outside the room. The Control group at this point received the intervention to guarantee that every participant had access to the same content

on the same day. The long-term follow up of the two groups is not an element in this research. The trial ended as planned as soon as the Control group finished attending the lecture.

Baseline data

In addition to the questions contained in the baseline and endline questionnaires (see Methodology) only a small set of characteristics were measured. Unfortunately, no socio-economic or inequality-related information was collected. Table 4 below summarises all data collected at baseline.

Table 4 - Baseline Characteristics			
	Overall	Treatment (T)	Control (C)
Average age	43.0 (6.76)	41.47 (7.07)	44.53 (6.31)
Average teaching experience ¹	12.5 (3.56)	11.53 (3.34)	13.47 (3.62)
Average sex (female = 1; male = 0)	0.73 (0.45)	0.6 (0.51)	0.87 (0.35)
Number of teachers	30	15	15

1. Teaching experience corresponds to secondary school only

It should be noticed that the table above does not show any statistical test related to the significance of the differences at baseline (ex: p-value). Such absence is due to a recommendation contained in CONSORT-SPI 2018 which says that

any difference at baseline happens by chance. Therefore it is illogical to provide what significance tests offer in this case. (Grant et al., 2018:12). What is considered relevant accordingly to CONSORT, is whether the size of the observed imbalances, for example, the 1.94 difference on the average teaching experience between the Treatment and Control groups, is of "prognostic importance" (Altman, 1985:126). In fact, according to Altman (1985), "imbalance in a baseline variable is only potentially important" when "that variable is related to the outcome variable" (p. 128). When we take the average teaching experience variable as an example, the question then becomes whether a 1.94-year difference (see Table 4, row 4) in this variable may affect, in practical terms, the comparability of the two groups. Altman (1985) says that to answer this question, "subjective assessment is preferable to significance testing, but requires prior knowledge of the prognostic importance of the variables being compared" (p. 130). Along these lines and sticking to the "Average teaching experience" variable example, we may leverage on, for example, the "five-stage theory" described by Berliner (2004), to identify the development stage in which the vast majority of this study's participants are very likely to be. With a high probability, a teacher who has 11 years of classroom experience is at the "Competent" stage, at least (Berliner, 2004:207). From this stage on teachers are said to be able to "make conscious choices about what they are going to do", "set priorities" and

"decide on plans" (Berliner, 2004:207). Saying it differently, from this moment on teachers can "determine what is and what is not important" (Berliner, 2004:207). Also, the theory says that after the Competent stage, the further development of the teacher brings primarily speed, fluidness and flexibility to his or her behaviour (Berliner, 2004:207). Leveraging on this theory, I consider it safe to claim that, above the 11-year threshold, it is unlikely that a 1.94-year difference in average classroom experience would significantly change the effect of a lecture on teachers' average self-efficacy for student-centeredness teaching, the outcome of interest. The intervention is frugal, and teachers' pedagogical orientations and choices are supposed to be already solidified at that stage.

A similar analysis can be undertaken about the other variables measured at baseline and shown in Table 4. Indeed, it looks even safer to subjectively accept, in what relates to self-efficacy for student-centred teaching", that those imbalances in "Average sex" (0.27) and "Average age" (3.03) have an even lower probability of harming the comparability of the two groups. One way or the other, it should be emphasised that differences at baseline, should we conclude that they have the potential to impact the "internal validity" of the RCT negatively, can be "controlled for" by quantitative techniques, such as multivariate analysis. (Duflo,

Glennerster, & Kremer, 2007:3949). For simplicity, these methods were not used in this study's case. Table 5 below shows measures of the outcome of interest at baseline. Both groups, Treatment and Control, answered the same questionnaire (see Methodology) before the Control group leaving lecture's room. Numbers in Table 5 represent average levels of self-efficacy, which were normalised to the interval [0,1]. The corresponding standard deviation is shown between parentheses.

Table 5 - Level of self-efficacy at baseline			
	Overall	Treatment (T)	Control (C)
	Mean	Mean	Mean
	(SD)	(SD)	(SD)
Helping students to identify strategies to support their own learning	0.695 (0.201)	0.693 (0.239)	0.697 (0.163)
Modelling strategies that students can use beyond the academic year	0.708 (0.198)	0.697 (0.232)	0.720 (0.166)
Providing opportunities for students to set goals	0.729 (0.199)	0.724 (0.225)	0.733 (0.176)
Helping students to recognise the value of the material	0.732 (0.179)	0.731 (0.216)	0.733 (0.140)
Providing a rationale to make academic tasks relevant	0.756 (0.203)	0.715 (0.243)	0.797 (0.152)

Creating opportunities for students to make choices about their learning	0.695 (0.175)	0.703 (0.203)	0.687 (0.149)
Helping students to monitor their performance appropriately	0.794 (0.166)	0.815 (0.200)	0.773 (0.128)
Assisting struggling students in identifying strategies to support their learning	0.700 (0.191)	0.714 (0.203)	0.687 (0.185)
Presenting content that students relate to other subjects	0.793 (0.152)	0.767 (0.187)	0.820 (0.108)
Creating authentic learning opportunities to make knowledge relevant	0.798 (0.166)	0.753 (0.202)	0.843 (0.108)
Recognizing students' thoughts in the explanation of learning tasks	0.812 (0.176)	0.777 (0.207)	0.847 (0.136)
Global (Mean of Averages)	0.747 (0.046)	0.735 (0.038)	0.758 (0.061)

Outcomes and estimation

As it was described in the Methodology section, while the Treatment group was attending the lecture, Control group participants were asked to fill the "endline questionnaire" to allow for the observation of the "counterfactual", that is, to

provide an estimation of what happens, "in expectation", when teachers are not exposed to the intervention (Duflo et al., 2007:3899). The Treatment group then finished attending the lecture and subsequently filled up the same questionnaire. Table 6 below presents the average differences of the type "endline - baseline", obtained from the answers given by the two groups, together with the standard deviation. The 95% confidence interval represents precision. The "Average Treatment Effect (ATE)" is estimated for each item and is represented both as a "mean difference (MD)" and a "standardised mean difference (SMD)" (Takeshima et al., 2014:1). P-values were calculated regarding the NULL hypothesis (ATE = 0).

Table 6 - Change in the level of self-efficacy (endline minus baseline)				
	Treatment (T)	Control (C)	ATE	p value
	Mean (95% CI) (SD)	Mean (95% CI) (SD)	MD (95% CI) ¹ SMD	(ATE = 0) ²
Helping students to identify strategies to support their own learning	0.104 (0.014, 0.194) (0.162)	0.043 (-0.033, 0.119) (0.137)	0.061 (-0.014, 0.135) 0.303 ³	.102

Modelling strategies that students can use beyond the academic year	0.100 (0.008, 0.192) (0.166)	0.047 (-0.042, 0.135) (0.160)	0.053 (-0.021, 0.128) 0.268	.146
Providing opportunities for students to set goals	0.053 (-0.045, 0.150) (0.176)	0.010 (-0.046, 0.066) (0.100)	0.043 (-0.060, 0.146) 0.211	.391
Helping students to recognise the value of the material	0.059 (-0.019, 0.138) (0.142)	0.017 (-0.042, 0.076) (0.106)	0.042 (-0.062, 0.147) 0.235	.396
Providing a rationale to make academic tasks relevant	0.103 (-0.000, 0.205) (0.185)	0.013 (-0.028, 0.054) (0.074)	0.090 (-0.039, 0.218) 0.443	.159
Creating opportunities for students to make choices about their learning	0.043 (-0.040, 0.127) (0.151)	0.080 (0.033, 0.127) (0.084)	-0.037 (-0.135, 0.062) -0.211	.438
Helping students to monitor their performance appropriately	-0.049 (-0.139, 0.042) (0.164)	0.030 (-0.009, 0.069) (0.070)	-0.079 (-0.175, 0.017) -0.476	.101
Assisting struggling students in identifying strategies to support their learning	0.099 (0.030, 0.168) (0.125)	0.060 (-0.002, 0.122) (0.112)	0.039 (-0.043, 0.121) 0.204	.321

Presenting content that students relate to other subjects	-0.014 (-0.085, 0.057) (0.128)	-0.041 (-0.109, 0.027) (0.123)	0.027 (-0.078, 0.133) 0.178	.587
Creating authentic learning opportunities to make knowledge relevant	0.001 (-0.069, 0.070) (0.126)	-0.073 (-0.153, 0.006) (0.144)	0.074 (-0.040, 0.188) 0.446	.187
Recognizing students' thoughts in the explanation of learning tasks	-0.033 (-0.108, 0.042) (0.136)	-0.020 (-0.083, 0.043) (0.115)	-0.013 (-0.102, 0.076) -0.074	.753
Global (Mean of Averages)	0.042 (0.004, 0.081) (0.058)	0.015 (-0.015, 0.045) (0.045)	0.027 (-0.007, 0.061) 0.148 ⁴	.104

- (1) Confidence intervals were computed using the "ci()" function which is part of the "gmodels" package in R.
- (2) P values were calculated using the R function "t.test()", and default parameters.
- (3) SMD numbers were calculated by dividing each MD by the standard deviation observed at baseline.
- (4) Contrary to the other items, this SMD was calculated by dividing the average ATE (MD) by the "average standard deviation" observed at baseline. I

Availability of trial data

Raw data collected during this experiment was made publicly available and can be downloaded from <https://doi.org/10.5287/bodleian:5Rjwgd8nV>.

Limitations

The most apparent limitation of the RCT described in this dissertation was the sample size, which was restricted by a “real world” constrain, as it sometimes happens with RCTs. In this study, the number of teachers invited by the Secretariat of Education circumscribed the number of participants to 32. As it was seen before, Treatment and Control groups ended up with 15 members each, which made it not possible the rejection of the NULL hypothesis ($ATE = 0$) with the probability that the industry normally accepts as a minimum: 0.80 (Cook & Hatala, 2015:74)). This insufficiency of “statistical power” (Cook & Hatala, 2015:74) happened even when the item’s estimated ATE was the largest in module: 0.476 standard deviations (see Table 6, row 8). In fact, a sample size calculation done with the help of the “pwr.p.test” R function indicated that 35 teachers (not only 15) would be necessary for each group to bring “enough” statistical power to the experiment when the effect size is this big (0.476 stdev), all other things being equal. The case for the smaller observed effects is even more unfavourable.

The issue mentioned above also relates to the magnitude of the p-values shown in Table 6. Commonly, only values $<.05$ are perceived as being indicators of statistical significance (Palesch, 2014:1). None of the p-values shown in Table 6 satisfies this criterion. While it is beyond the scope of this

dissertation to detail the problems associated with equating “ $p < .05$ with statistical significance” (Palesch, 2014:2), the cautious reader should have in mind that a p-value is also a function of the sample size. Saying it differently, “you can make the p-value as small as you can afford” (Demidenko, 2016:33), that is, “with a large enough sample, n , the NULL hypothesis will always be rejected” (Demidenko, 2016:33). Along these lines, it makes little sense to conclude that merely because p-values are greater than .05, the interpretation of this experiments’ results is of no value. In this respect, the American Statistical Association (ASA) recently issued a statement which says: “a relatively large p-value does not imply evidence in favour of the NULL hypothesis” (Wasserstein & Lazar, 2016:132).

Also, the p-values shown in Table 6 do provide an intuition regarding the likelihood of observing an ATE “as extreme or more extreme” than the one estimated for that type of self-efficacy, in a situation where the “real” ATE is inexistent (Palesch, 2014:2). In this sense, given that the sample size was kept fixed amongst the questions of the questionnaires, the lower the p-value, the more unlikely it would be the estimation of the corresponding ATE if the intervention (treatment) has caused no effect. This notion, combined with the actual size of the estimated ATE in standard deviations, allows for the identification of the types of self-efficacy to

which the intervention was probably able to produce the most "salient" effects. Attempts to interpret experiments' results may then prioritise those cases.

Generalisability

The generalisability of this experiments' results depends upon what Cartwright & Hardie (2017) call "support factors" (p. 273). In this regard, it should be understood that the structure of the experiment described herein was designed to be imbedded into an existent teacher education program so to provide evidence of its elements' effectiveness. Therefore, it seems safe to assume that the support factors which made it possible for this intervention to produce an effect, will also be in place next time the same intervention is executed within the same program. Indeed, the more the experiment is repeated with different groups of teachers, the higher our ability to predict what is going to happen when we intervene.

Interpretation

Cohen (1990) famously stated that "the primary product of a research inquiry is one or more measures of effect size, not p-values" (p. 1310). Also, he said that "a successful piece of research doesn't conclusively settle an issue, it just makes some theoretical proposition to some degree more likely" (p.

1310). Having this position in mind, and given the fact that all estimated ATEs shown in Table 5 failed to cross the “ $p < .05$ cliff” gap (Cohen, 1990:1311), I considered it more adequate to offer here an attempt to interpret only the result which surprised me the most. My lecture on OKR ended up reducing by almost 0.5 standard deviations teachers’ self-efficacy for helping students to monitor their performance (see Table 6, row 9). This item’s ATE (-0.476) was the largest in magnitude and the one with the smallest p-value (.101).

At this point, I consider it easy for the diligent reader to imagine why I found it surprising the result mentioned above. Indeed, goal-setting in the classroom is commonly associated with students, in different degrees, self-regulating their learning (Pintrich, 1995:6; Zimmerman, 2002:65; Zimmerman & Kitsantas, 2014:145). At the same time, the ability to self-monitor one’s performance is considered to be a foundational process in self-regulation (Huh & Reigeluth, 2017:199). Accordingly to Zimmerman & Kitsantas (2014), goals “enable learners to create self-oriented feedback loops to monitor their effectiveness and to adapt their functioning” (p. 145). Because of the above, I was convinced at first that teaching teachers the fundamentals of a modern goal-setting framework would raise, not reduce, their self-efficacy for helping students to monitor their performance.

After reflecting upon a set of potential causes of this unexpected result, and assuming that data was rightly manipulated, I came up with an explanation which is at this point purely speculative since the estimated ATE was not statistically significant. I believe that it is worthwhile further investigating, in a different opportunity, whether this large negative ATE was the result of the lecture on OKR, the intervention, functioning as a "wake up call" regarding what those teachers were doing to help students to evaluate their performance. Saying it differently, I guess those teachers did not know how far they were from being able to foster more self-regulation in their students. As a possible indication of this perception-reality gap, we have the fact that this type of self-efficacy was the one to present the highest level at baseline (see Table 5, row 9). In this regard, it seems to me that my colleagues got frustrated when it became evident that they were not as capable as they thought they were. I considered it possible that this frustration might have over-reduced their self-efficacy at endline. While it is beyond the scope of this dissertation to further explore this phenomenon, I believe that this decrease in self-efficacy just after a teacher education intervention, if verified to happen in other occasions, is something that teacher educators should understand in greater detail.

Finally, it must be emphasised that the proposition described above is just one possibility amongst many. The repetition of this experiment is perhaps the right way to make it more likely or refute it entirely.

A narrative case study of OKR in the classroom

Flyvbjerg (2006) says that accordingly to the "conventional wisdom of case-study research" this methodology is only useful at "the preliminary stages of an investigation" (p. 220). At the heart of such notion, he states, there is a common misunderstanding which says that "general, theoretical (context-independent) knowledge is more valuable than concrete, practical (context-dependent) knowledge" (p. 221). Notwithstanding this notion, it is widely known that expert teachers, for example, operate "on the basis of intimate knowledge of several thousand concrete cases" (ibid, p.222). The knowledge in possession of teachers becomes so intimate as time goes by that ends up amalgamating with practice and forming what some authors call "gestalts" (Korthagen & Kessels, 1999:13). These elements are said to influence teachers' "split-second behaviour" (Korthagen & Kessels, 1999:13) and can hardly be changed by the simple acquisition of context-independent knowledge. Flyvbjerg (2006), summarises his argument in favour of context-dependent knowledge by saying

that "it is only because of experience with cases that one can at all move from being a beginner to being an expert" (p. 222).

Inspired by the above, my intent in the lines bellow is to offer a nuanced view of a real-life situation, in this case, the first attempt to implement OKR in a classroom. I shall provide as much wealth of details as possible so to make it easier for the reader to "identify with the experience" of the teacher who worked with me (Brandell & Varkas, 2001:294) and also to facilitate the life of those who decide to reproduce the steps described herein. The guidelines presented in Baxter & Jack (2008) and (Brandell & Varkas, 2001) will be followed whenever applicable.

Context

Karen (see Methodology), works as a teacher both for the State and for the City of Rio de Janeiro. In Brazil, the government is supposed to offer 9 (nine) years of compulsory schooling. Municipalities, not the states, are the ones responsible for building and running primary and secondary state-owned schools (Louzano, 2007:11). It happens that due to a variety of circumstances, it is quite common to find secondary or even primary schools being run by states. This situation makes it frequently that teachers end up working for more than one instance of government. *Karen* is a seasoned

teacher with more than 15 years of teaching experience working for the City, and more than nine years of experience working for the State. Also, an interesting aspect related to Karen is the fact that she was recently promoted to become a teacher educator, without abandoning her role as a teacher, at the State of Rio. This new condition of hers, in addition to the extra work, created unexpected and context-related difficulties, as it can be seen in the excerpt below.

I think that I don't want to be a teacher educator anymore. The way the Secretariat is conducting things makes it very dangerous the job. Picture the following situation. She asked me and others to go to a distant school to explain our newest initiative to teachers. The problem is that this school is in the middle of a "comunidade." I said: 'I am not going. Local gang members know those teachers. They do not know me. I don't want to get shot'.

The quotation above illustrates a reality that is very particular to the City of Rio and perhaps to other underdeveloped large cities across the globe. "Comunidade" means literally, "community" and is a common euphemism for "hazardous poor area." What is relevant in the context of this dissertation is the perception that the OKR pilot was taking place within a rather precarious educational system. Deole (2017) says that violence presents "a negative externality on educational outcomes in the Brazilian system" (p. 307). The refusal of Karen described above is almost certainly an element of this externality.

After Karen told me about the event above, I questioned whether she believed that the use of OKR was something we should care about, considering those more critical issues. She answered:

I think so, and I will tell you why. I feel that these kids are accustomed to receiving almost nothing from their teachers. At the same time, we teachers are not motivated to do more. Maybe we should all start talking about having new goals and see if this situation changes.

Indeed, I had already detected this tacit agreement between demotivated teachers and disengaged students in a previous assignment. The case for English as a Foreign Language (EFL), the subject taught by Karen, is particularly problematic. Brazilians strongly believe that "they do not learn English at regular schools" (Bambirra, 2017:216). Therefore, the situation can be summarised as follows: demotivated teachers teaching students who *a priori* consider it useless the time spent in English classes. Regarding this complex scenario, recent research shows that "emotional and (meta)cognitive" factors are the most predominant elements which influence Brazilian student's engagement in learning English (Bambirra, 2017:234). In this sense, given that the pursuit and attainment of goals are frequently associated with "cognitive and affective-based processes that enhance positive achievement activity" (Lemos, 1996:152) and based on Karen's response to

my inquiry, I suspect that even before we have ignited the OKR pilot, she already had an intuition that the use of OKR as a “language” to talk about performance would make it possible a new type of agreement, this time explicit, between teachers and students.

Phase #1 – Set

OKR can be understood as being a continuous cycle encompassing three phases: Set, Align and Achieve. The Set phase corresponds to the moment when it is defined, in writing, the Objectives and Key Results which will be pursued by an individual or group during the cycle (Doerr, 2018:37).

Karen and I started to have weekly Skype sessions (see Methodology) in May 2018. At first, it was necessary for us to agree upon a certain OKR cadence, that is, the lapse of time through which each OKR cycle (Set, Align and Achieve) was supposed to be completed. My first suggestion to Karen was that she should organise monthly 20-minute goal conferences (Gaa, 1973:22) with her students. She replied by stating:

I don't think I will be able to cope with the extra workload which is required to meet individually with dozens of students every month. You should remember that I teach in two different schools and that I have at least three classes in each school. Why not do it once in each bimester? Our objectives here are always defined in a bimestrial basis.

Teachers often claim that they “have little time to talk to pupils” or that they are “under so much pressure to reach standards.” These complaints are said to be part of an international pattern of discontent (Scott, Stone, & Dinham, 2001:12). In this sense, it seemed logical to me that *Karen* would have second thoughts about implementing OKR, the moment she knew more about the extra work required. What I had not anticipated was the fact that even the idea of having a 20-minute encounter with each student once a month would be considered an almost insurmountable hurdle in her case. Luckily, the second thing she said, that is, the suggestion to run OKR with the same periodicity by which students are already evaluated, made sense to me. Doerr (2018) says that the “best OKR cadence is the one that fits the context and culture” (p.36). Therefore, it seemed reasonably right at that point the idea to synchronise students’ goal-setting with the formal evaluation process.

After having agreed upon a 2-month cadence, *Karen* and I had to face the challenge to transpose to the educational context the mixed “bottom-up and top-down” goal definition process (Zhou & He, 2018:321) which constitutes one of the strengths of the OKR method. Contrary to the Management by Objectives (MBO) framework, for example, OKR is a system where objectives are primarily set from the bottom up (Radonic, 2017:29). Such aspect means that *Karen* and I wanted students,

who had never heard about these things before, to be able to define their objectives for the bimester. It should not be forgotten that we were piloting OKR in secondary school. That means dealing with 11-16-year-old kids. It ended up taking us three sessions of trial and error until Karen was comfortable enough to proclaim mission accomplished. The lines below describe what happened.

Objectives in OKR are defined accordingly to a unique set of guidelines. Among other aspects, goals are supposed to be inspiring and engaging (Radonic, 2017:30). We talked about this concept in one of our first sessions, and there she went Karen to ignite the process with students the next day. The short dialogue below was extracted from a video file sent to me by Karen. These were 12-13-year pre-adolescents.

Karen: *Why are you here?*

Student #1: *Because we must.*

Karen: *I am not joking. I will rephrase. Why studying English?*

Student #2: *To be able to talk to people who don't speak Portuguese.*

Karen: *Only because of that? What happens when everybody around you speaks Portuguese? Is English less important in this case?*

Student #3 *Instagram.*

Student #4 *Not me. I only use Portuguese on Instagram.*

It is unnecessary to say that *Karen* did not go far with her first attempt to make students reflect on their objectives. She almost lost the control of the classroom as students kept disagreeing with each other.

As a short and parallel comment, it must be understood that promoting the formation of non-cognitive skills such as the ability to set and achieve goals, is a difficult task. Research on "Social and Emotional Learning (SEL)", a concept which relates, among other aspects, to teachers fostering goal-setting competencies (Osher et al., 2016:646), indicates that teachers frequently claim to need "strong support from district and school leaders" so to "effectively implement and promote SEL skills" (Schonert-Reichl, 2017:140). In this sense, one should expect that an untrained teacher such as *Karen* would need a lot of help to be able to implement OKR in a classroom, especially when there is no practical information from where to draw.

Karen and I spent two additional sessions discussing the best approach to apply OKR's goal definition guidelines. I insisted to her that the determination of goals was supposed to be primarily bottom-up otherwise we would be piloting something else, not OKR. She was concerned that if students were left alone their goals would end up not being aligned with the curriculum or would be everything but inspiring, engaging and memorable. At some point *Karen* said:

What if I give them a list of options? For example, I can meet with each of them individually and ask the student to pick a goal from a pre-defined list.

It was not clear to me whether providing a list of goal options to kids would negatively affect the framework's impact. At the same time, I was aware of the OKR concept which considers managers' OKRs a "vision" or "principle" that needs to be communicated to the rest of the company and ends up influencing the formation of the other OKRs (Doerr, 2018:35). In this sense, OKR is not just "bottom-up" after all. I said to myself that maybe teachers are the ones who should play, in this respect, the role of the managers, perhaps just until students get accustomed to the framework. With that in mind, I agreed to give it a try the list idea. Karen then suggested the following objectives to students:

1 - I will translate the lyrics of a song that I love

2 - I will translate an interview given by a Brazilian football player who plays in the Premier League

3 - I will create subtitles for a 60-second scene of a Netflix series that I enjoy watching

The proposed solution worked. Each student picked one of the objectives shown in the list above, and then Karen moved on to presenting the concept of each goal having one or two

Key Results associated with it. The short dialogue below illustrates what happened.

Karen: Now that you all have objectives...how are we going to know whether you have achieved what you planned?

Student #1: Will you give me an A if I achieve 100% of my goal?*

Karen: Not at all. Your grades will be unaffected by the achievement. However, I will evaluate your effort in participating. Your enthusiasm will either raise or reduce your final grade. Going back to my question, how are we going to measure your achievement? We must define a criterion. It should have a number in it.

Karen had already anticipated that students, at some point, would start talking about linking grades and marks to the achievement of OKRs. However, within the organisational field OKRs are typically not associated with financial incentives or other "extrinsic rewards" (Doerr, 2018:159; Zhou & He, 2018:321). Following the same idea, Karen and I decided to make it detached from marks and grades the achievement of OKRs. By doing so, I knew that we were emphasising the setting of "mastery goals" over "performance goals" (Gonçalves, Niemivirta, & Lemos, 2017:150; Lemos, 1996:152; Moeller, Theiler, & Wu, 2012:154). This decision appeared to be adequate since this type of goal is considered to cause a more significant positive impact on the self-regulation of learning (Mykkänen, Perry, & Järvelä, 2015:3).

Going back to the task of defining Key Results, Karen and I faced, in this case, the same "bottom-up versus top-down" dilemma. This time we were quick to decide that Karen should interfere so to speed up the whole process. As time goes by, we expect the amount of teacher interference to stabilise at a much lower level. Here is how an OKR ended up looking after Karen finished defining the Key Results together with her students:

*1 - I will translate the lyrics of a song that I love
(Objective)*

as measured by the following process:

I will sing along the Portuguese translation while the English version is being played..

My teacher will interrupt me five times to ask the meaning of a word or expression. I can make up to 1 mistake. (KR 1)

An adult family member will evaluate my translation and will write an email or note to my teacher explaining his or her criteria. 80% is the minimum score that I am supposed to receive. (KR 2)

It should be noticed that the Key Results in the example above combined teacher-assessment with the assessment made by a family member. The involvement of caregivers with the OKR pilot was Karen's idea. Here is what she said to justify such initiative:

One of our objectives for the present year is to make it more fluid the family-school connection. The moment you have shown me the OKR framework I realised that it

could be used to make it more dynamic this relationship.

This possibility looked promising to me. I was fully aware that quite often families from poor social classes are considered to be “not interested” and “absent from meetings and the monitoring of their children’s education” (Hellen, Marcondes, Regina, & Lucato, 2012:93). Intuitively, I have always refused to endorse such a negative conception. In addition, now that the digital age is making it widely available the information needed for students to self-direct their own learning (Lee, Osop, Hoe-Lian Goh, & Kelni, 2017:612), why not leveraging on the same capabilities to engage their families with the assessment for learning process (Rutland & Hall, 2013:116-117)?

After going through the events described in this section, Karen and I finally managed to have the first group of students fully involved with an OKR cycle. Then it was time for us to reflect on the alignment phase.

Phase #2 – Align

In the organisational field, recently created OKRs need to be aligned before the individual or group starts putting effort to achieve what was defined (Castro, 2018:1). Alignment within

a business means discussing the OKR with one's boss to make sure that the OKR presents all the desirable characteristics (memorable, bold, measurable etc.). Also, in the case of OKRs depending on other OKRs these interdependences need to be mapped and adequately addressed before hand. In a real business scenario, OKRs should be aligned not only with the overall strategy but also with other OKRs (Radonic, 2017:31; Zhou & He, 2018:222). In this sense, it can be said that an OKR may be freely drafted, but its adoption depends upon the OKR going through an alignment process in which the management of the company has to either approve the OKR or ask the individual to rewrite it.

Karen and I were fully aware that her approval was an element of the Align phase and that it was supposed to happen after students have by themselves drafted the OKRs. Imagine, for example, if *Karen* did not want, for some reason, two students to use the same song in their OKRs. In this case, the Align phase would allow for this type of orchestration. However, as it was described herein in a different section, the low "maturity level" of *Karen's* students in regards to goal-setting made *Karen* interfere during the definition of OKRs. Because of this, the Allign phase was absorbed by the Set phase in this first pilot. *Karen's* students' OKRs were born already alligned.

I asked Karen what would be the best way to coordinate the Set and Align phases when students become capable of drafting their OKRs without her help. She answered:

I think that perhaps the only way possible is to ask students to bring already drafted OKRs to the goal conference. I believe that during the time they will be able to check the next chapters of the book and imagine what they want to accomplish in the new bimester. Maybe I am asking too much.

I found it intriguing this idea to ask students to visit the textbook, reflect on the topics which are planned to be discussed in the next bimester, and draft an "OKR proposal". Such an approach seems to be aligned with the literature on flipped education (Schwarzenberg, Navon, Nussbaum, Pérez-Sanagustín, & Caballero, 2018:238) and will be further investigated in future research.

Phase #3 – Achieve

The Achieve phase corresponds to students doing what they need to do to accomplish what was specified by themselves in their OKRs. Karen found a significant level of heterogeneity in her students' engagement during this phase.

Some students used to be very demotivated before the pilot. Now they have completely changed, and their OKRs appear to be the only thing they want to talk. Isn't it wonderful? On the other hand, I have one or two cases where the student used to be very participative and now is not demonstrating the same

enthusiasm. Don't know whether they are doing something.

I asked *Karen* what method she was using to track each student's progress and she said:

I don't use any technique. They tell me about their OKRs. Generally at the beginning of the lesson.

I asked *Karen* to promote weekly "OKR checkins" (Radonic, 2017:33) so to track everybody's progress without depending on students' initiative to disclose information. She said that she did not have space in the agenda to accommodate a ceremony of that type. I then suggested the use of a "check-in questionnaire" to be filled by each student at the beginning of the lesson. The questionnaire might contain the following two questions:

I am _ % satisfied with the progress of my OKRs
It is _ % urgent that I discuss some OKR's issues with the teacher

Table 7 - Check-in questionnaire

Karen mentioned that she used the check-in questionnaire described above just once. She said that this was so because we were already approaching the end of the pilot when it was developed. She once again complained that even with the use of

the questionnaire she did not have enough space in her agenda to be able to sit down and discuss students' issues.

I would love to dedicate myself to discussing objectives fulltime. However, I am supposed to teach the curriculum. You know how things are in this city. The other week this school was closed due to the eminence of violence in the neighborhood. It was one day less to do everything that I have to do.

Due to what Karen said to me, I started to ask myself whether it would be possible for a teacher who has dozens of students under her responsibility, to accurately track, on a weekly basis, the evolution of each OKR. A classroom is not equivalent to a company's department. People in a department work together for eight hours and pursue together more or less the same objective. Cerimonies can be quickly organised in such context. Students in a classroom, on the other hand, go from subject to subject during a 5-hour period. A teacher like Karen stays in front of her students only two hours per week. Will this limited availability of time allow for the promotion of weekly checkins? Here is what Karen had to say in this respect:

Now that you asked, I don't think so. It was ok helping students to define the OKRs. Can the follow up be done by a different person?

In the end, the vast majority of students achieved the OKRs. The follow up, Karen and I admit, was far from being perfect. We ended up concluding that there should be a dedicate

person, in each school, responsible for conducting the weekly checkins with students. This person could be named the OKR "Evangelist" or "Shepherd" (Doerr, 2018:70) and should be adequately trained to be able to deal with OKRs linked to all subjects in the curriculum.

Limitations

The OKR pilot which I described above using the "narrative case study" methodology aimed to identify a "manner of use" for the OKR framework inside a classroom. Given that there was no practical information to draw from, *Karen* and I ended up making some mistakes, the most serious of which is the nature of the OKRs that we induced her students to define. We made the common mistake to think about "outputs" (activities) not "outcomes" (Mosse & Sontheimer, 1996:9). In other words, we allowed an "activity" to be treated as the Objective in an OKR. Saying it differently, instead of making students to reflect upon memorable and bold results, such as, for example, "I will understand English so well that I will be able to understand the songs that I love", we made students consider as the Objective the activity which corresponds to successfully translating a single song. This activity was supposed to be regarded as part of a Key Result which could be then associated with the bold and memorable objective indicated above. *Karen*

and I have recently started a new OKR cycle in which we did not make the same mistake.

Another limitation of the case study described in this section is the fact that it corresponded to the experience of a single teacher. In this sense, it should be emphasised that by no means I find myself in a position to make general claims regarding the utility of the OKR framework as a tool capable of fostering engagement and motivation in secondary classrooms.

The effect of OKR on *Karen*, though, seems to be more evident. The adoption of the framework induced her to reflect upon her objectives. While it is unclear whether other teachers would be impacted the same way, *Karen* finished the pilot with a different view of her professionalism. She said that all of a sudden, being a teacher became much more than merely knowing the best way to teach something (Shulman, 1986:9). The act of teaching was transformed into, in her words, "knowing how to command a group of kids who want to learn and achieve something". I recognise that the proximal relationship which I have maintained with her during the pilot might have substantially changed her perception of things. However, it is impossible not to highlight that this new understanding of her professionalism is in line with what is described today by some authors when they talk about leadership inside the

classroom (Katyal & Evers, 2004; Öqvist & Malmström, 2016, 2018).

Finally, I must say that this research on OKR in the classroom is just beginning. If things go well I will be able to deploy the framework in a more significant number of schools soon. Such a possibility is, in fact, in one of my OKRs.

Conclusion and Implications

The effect of a lecture on goal-setting

The possibility that goal-setting, and more specifically, the OKR framework could be used by teachers to change behaviour so to engage students, made me interested in investigating whether a teacher education intervention on goal-setting would affect teachers' set of beliefs and practice. In my view, this would be an exciting starting point for an OKR implementation. Because of this, the quantitative phase of this research aimed to understand whether a typical teacher education session on OKR, that is, a lecture, would positively impact the self-efficacy of teachers regarding student-centred teaching. Also, this phase intended to evaluate the specifics of using randomised experiments to measure the effectiveness of professional development programs.

Findings indicated that a 50-minute lecture on goal-setting was, in that occasion and in-light of those specific conditions, capable of raising teachers' self-efficacy for student-centred teaching. The number of participants (sample size) was not large enough to allow for sufficient experiment's power ($> 80\%$) (Cook & Hatala, 2015:74) so to detect even a mid-sized difference (SMD 0.5 standard deviations) (Cook & Hatala, 2015:73) with sufficient statistical significance. Therefore, it is not possible to say at this point whether the lecture "caused" the detected difference shown in the findings or it just happened by chance. However, as it was discussed in the methodology section, the conciseness of the RCT method used in this study allows for its easy repetition with different cohorts so to increase experiment's statistical power through meta-analysis (Cohn & Decker, 2003:243).

In this sense, the most crucial aspect revealed by the quantitative phase is the applicability of an RCT, a method that is frequently considered controversial in education and teacher education circles, to produce evidence on the effectiveness of an intervention as frugal as a single lecture. In reality, most of the issues which are commonly associated to an RCT, such as, for example, the difficulty in generalising results, can be attenuated, "in expectation" (Deaton & Cartwright, 2018:5), when the trial can be repeated an arbitrary number of times. In theory, for example, the larger

the sample size or the number of trials, the higher your ability to detect the average effect caused by an intervention. In this regard, the most fundamental question related to the use of an RCT to inform the evaluation and perhaps re-design of teacher education programs is whether and by which cost the experiment can be repeated so to increase "the warrant for an effectiveness prediction" (Cartwright, 2013:105).

Having said that, it should be noticed that this study has identified straightforward, practical ideas, such as, the creation of equivalent but apparently different questionnaires to measure the outcome of the intervention at different moments, which can make it possible, if not easy, for an RCT to be used in a variety of real professional development situations. These concepts can be adopted and further developed by anyone interested in investigating the effectiveness of a teacher education intervention via the use of a "micro-RCT".

Due to all the above, the quantitative research question ("What is the effect of a teacher education intervention on goal-setting on teacher's self-efficacy for student-centred teaching?") could not be answered. Yes, it looks like that the lecture on OKR, which aimed to increase teachers' self-efficacy for student-centred teaching, "causes the targeted effect in some individuals in some settings" (Cartwright, 2013:98), at least immediately after they have attended the lecture. However, despite all other promising aspects connected to the

execution of the RCT and described in the lines above, the sample size did not allow for any claim stronger than that.

The practicalities of using OKR in the classroom

After conducting a randomised experiment which has shown that a lecture on OKR has the potential to affect teachers' set of beliefs, I aimed to investigate whether it would be practical for a teacher with no previous experience with goal-setting, to adopt the OKR framework in her classroom.

Given that OKR, to the best of my knowledge, had never been systematically used in a classroom before, this investigation required the adaptation of the framework to the educational context. In this sense, some of my interactions with the teacher who I worked with, had me playing the role of a teacher educator who was helping her to adapt and implement something that was new to both of us. Besides that, I also interacted with this teacher in the capacity of an educational researcher who was trying to learn about the main issues related to the implementation of the framework, as perceived by a teacher. This duality made it more productive the whole experience for me as a teacher educator. At the same time, it made it more difficult the separation of my own voice from hers when I was looking for an answer for the research question associated with this qualitative phase ("What do teachers

perceive as being the key issues and challenges for teachers with no experience with goal-setting to learn and efficiently apply the OKR framework?“).

Overall, the findings revealed that the OKR implementation was considered a promising methodology by the teacher involved with the pilot. Three points of concern ended up being salient in her discourse: (i) The extra time and effort required by the preparation and execution of monthly individual goal conferences; (ii) Students needing help to be able to suggest “bold” OKRs linked to curriculum items continually; (iii) The need for a designated person to be responsible for conducting weekly OKR check-in ceremonies (follow-up) with students;

Despite those issues, the teacher said that OKR helped her to understand the learning stage of each student correctly, and, more importantly, allowed her to provide feedback more efficiently. The teacher strongly emphasised the notion that, with OKR, formative feedback comes more naturally as part of a new, formal, structured and “detached from marks and grades” in-school conversation about performance. She has perceived this new conversation as something able to prompt students to develop a more sophisticated self-reflection on learning.

Finally, it should be mentioned that teaching a teacher how to use OKR was more straightforward than expected during this research. Perhaps because of her condition as a professional who is part of an organisation (school/system),

the fundamental notion that the systematic pursuit of goals helps achievement, seemed to make a lot of sense to her since the very beginning. The three main issues detected were anticipated by the teacher during our early discussions and adequately circumnavigated. At the same time, the positive aspects of the method (being bottom-up; requiring quantitative criteria; valuing boldness) were considered by her of great importance and capable of increasing student autonomy and engagement.

Non-cognitive skills and OKR

This research was ignited by in-school observations which had shown, in my context, students not demonstrating an appropriate level of engagement in learning and teachers not paying the necessary attention or just not knowing what to do to engage them. The possibility that a portion of this perceived reduction of student engagement was merely a function of students increasingly using the Internet for autonomous and "teacher-less" learning (Katyal, 2010:273) was an initial hypothesis in this research. According to this original view, the engagement issue which I had observed was being caused by technology-driven changes which affect countries in an almost equal way. These changes would end up enforcing teachers to adjust by taking "conscious steps so that they are clearly able to influence and promote school-based learning" (Katyal

& Evers, 2004:372), that is, would induce them to assume a role which is described by some authors as the role of a "developmental leader" (Öqvist & Malmström, 2018:156). Motivated by this notion, the bulk of the study corresponded to experimenting with a type of leadership behaviour (goal-setting) that teachers may adopt to affect students' motivation and engagement in present times.

As the research progressed, though, I started to have the impression that, in addition to potentially being caused by a general and Internet-related trend towards "teacher-less" learning, the relative lack of student engagement which I have observed was also a function of something else. Indeed, a very recent econometric study shows that non-cognitive skills, mainly those related to motivation and engagement, may partially explain the relationship between economic growth and PISA test scores (Balart, Oosterveen, & Webbink, 2018:146). According to this study, for any given level of economic development, treated in the article as a proxy for the quality of the local educational system, the variation in PISA scores amongst similarly developed countries can be explained by differences in non-cognitive factors, such as motivation, amid students. Because of this, the lack of student engagement which I have observed, may not be only a function of a global trend towards "out-of-school learning" (Katyal, 2010:276), given that such phenomenon would equally affect countries and would

not be able to explain the variation in PISA scores. There should be something else, perhaps culture-related (Balart et al., 2018:146), affecting engagement in learning and causing a reduction in academic attainment in Brazil.

In this sense, in the case context-related factors are verified to be harming the formation of non-cognitive skills which may be, in turn, impacting attainment in high-stakes exams such as the PISA, then there should be a stimulus for the local continuation of the research described in this dissertation, since goal-setting is considered a foundational process in student motivation and engagement. Indeed, it was salient in the discourse of the teacher who worked with the OKR pilot that some disadvantaged students who used to adopt a "couldn't care less" attitude, got involved with the OKR process and immediately started to behave differently. While it is unclear whether such improvement will translate itself into attainment in the future, the perception of the teacher is that the use of OKR might have unlocked a substantial amount of academic potential.

Implications and next steps

This research was about helping teachers to increase their students' engagement in learning. It investigated the use, in the context of teacher education, of a modern goal-setting

framework, the OKR. The next steps to take following this study will aim to solidify a "manner of use" for the OKR framework in the educational context. The pioneering aspect of such work should not be taken for granted.

The teacher who participated in the pilot and one of her colleagues are already starting OKR with a new group of students. Also, an authorisation for a more significant pilot next year is expected from the Secretariat of Education. Data from these further investigations will support an innovative stream of research, particularly concerning correlating OKR achievement with academic attainment.

At the methodological side, the design of "micro-RCTs" to be used by researchers and practitioners to produce evidence on the effectiveness of short term interventions has also potential to create innovation with practical consequences in the field of teacher education.

Finally, the effectiveness of the OKR method, as measured regarding academic achievement in high-stakes exams, has the potential to reveal the extent to which non-cognitive skills may represent a competitive advantage in a world where subject content and PCK are more available than ever before.

Bibliography

- Aghera, A., Emery, M., Bounds, R., Bush, C., Stansfield, B., Gillett, B., & Santen, S. (2018). A Randomized Trial of SMART Goal Enhanced Debriefing after Simulation to Promote Educational Actions. *Western Journal of Emergency Medicine, 19*(January), 112-120.
<https://doi.org/10.5811/westjem.2017.11.36524>
- Altman, D. G. (1985). Comparability of Randomised Groups. *Journal of the Royal Statistical Society. Series D (The Statistician), 34*(1), 125-136.
- Ames, C. (1992). Classrooms: Goals, Structures and Student Motivation. *Journal of Educational Psychology, 84*(3), 261-271.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260-267. <https://doi.org/10.1037/0022-0663.80.3.260>
- Balart, P., Oosterveen, M., & Webbink, D. (2018). Test scores, noncognitive skills and economic growth. *Economics of Education Review, 63*(October 2017), 134-153.
<https://doi.org/10.1016/j.econedurev.2017.12.004>
- Bambirra, R. (2017). Motivation To Learn English As a Foreign Language in Brazil - Giving Voice To a Group of Students

At a Public Secondary School. *Linguagem Em Discurso*, 17(2), 215-236. <https://doi.org/10.1590/1982-4017-170204-5316>

Baxter, P., & Jack, S. (2008). The Qualitative Report Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544-559. <https://doi.org/citeulike-article-id:6670384>

Berliner, D. C. (2004). Describing the Behavior and Documenting the Accomplishments of Expert Teachers. *Bulletin of Science, Technology & Society*, 24(3), 200-212. <https://doi.org/10.1177/0270467604265535>

Berson, Y., Halevy, N., Shamir, B., & Erez, M. (2015). Leading from different psychological distances : A construal-level perspective on vision communication , goal setting , and follower motivation. *The Leadership Quarterly*, 26(2), 143-155. <https://doi.org/10.1016/j.leaqua.2014.07.011>

Brandell, J. R., & Varkas, T. (2001). Narrative Case Studies. In B. A. Thyer (Ed.), *The Handbook of Social Work Research Methods* (pp. 294-307).

British Educational Research Association. (2011). Ethical guidelines for educational research. *British Educational Research Association*, 1-11. <https://doi.org/978-0-946671->

Bruns, B., Costa, L., & Cunha, N. (2018). Through the looking glass: Can classroom observation and coaching improve teacher performance in Brazil? *Economics of Education Review*, 64(June 2017), 214-250.

<https://doi.org/10.1016/j.econedurev.2018.03.003>

Bullock, K., & Wikeley, F. (1999). Improving learning in year 9: Making use of personal learning plans. *Educational Studies*, 25(1), 19-33.

<https://doi.org/10.1080/03055699997945>

Cartwright, N. (2011). A philosopher's view of the long road from RCTs to effectiveness. *The Lancet*, 377(9775), 1400-1401. [https://doi.org/10.1016/S0140-6736\(11\)60563-1](https://doi.org/10.1016/S0140-6736(11)60563-1)

Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn't always travel. *Evidence and Policy*, 9(1), 97-112. <https://doi.org/10.1332/174426413X662581>

Cartwright, N., & Hardie, J. (2017). Predicting What Will Happen When You Intervene. *Clinical Social Work Journal*, 45(3), 270-279. <https://doi.org/10.1007/s10615-016-0615-0>

Castro, F. (2018). The OKR Cycle: 3 Steps for OKR Success. Retrieved from <https://felipecastro.com/en/okr-cycle/>

Childs, A., & McNicholl, J. (2007). Investigating the relationship between subject content knowledge and pedagogical practice through the analysis of classroom

discourse. *International Journal of Science Education*, 29(13), 1629-1653.

<https://doi.org/10.1080/09500690601180817>

Cohen, J. (1990a). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312.

<https://doi.org/10.1037/0003-066X.45.12.1304>

Cohen, J. (1990b). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312.

<https://doi.org/10.1037/0003-066X.45.12.1304>

Cohn, L. D., & Decker, B. J. (2003). How Meta-Analysis Increases Statistical Power. *Psychological Methods*, 8(3), 243-253. <https://doi.org/10.1037/1082-989X.8.3.243>

Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education: Theory and Practice*, 20(1), 73-83.

<https://doi.org/10.1007/s10459-014-9509-5>

Day, T., & Tosey, P. (2011). Beyond SMART? A new framework for goal setting. *Curriculum Journal*, 22(4), 515-534.

<https://doi.org/10.1080/09585176.2011.627213>

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine*, 210(October 2017), 2-21.

<https://doi.org/10.1016/j.socscimed.2017.12.005>

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology, 49*(1), 14-23.

<https://doi.org/10.1037/0708-5591.49.1.14>

Deci, E. L., & Ryan, R. M. (2012). *Motivation, Personality, and Development Within Embedded Social Contexts: An Overview of Self-Determination Theory. The Oxford Handbook of Human Motivation.*

<https://doi.org/10.1093/oxfordhb/9780195399820.013.0006>

Demidenko, E. (2016). The p-Value You Can't Buy. *American Statistician, 70*(1), 33-38.

<https://doi.org/10.1080/00031305.2015.1069760>

Deole, S. S. (2017). Human Capital Consequences of Violence in Schools: Estimating the Impact of Violence in Schools on Education Outcomes in Brazil. *SSRN Electronic Journal, 287-310*. <https://doi.org/10.2139/ssrn.2925848>

Doerr, J. (2018). *Measure What Matters*. New York: Portfolio/Penguin.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write managements's goals and objectives. *Management Review*.

<https://doi.org/10.1177/004057368303900411>

Duflo, E., Glennerster, R., & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics, 4*(07), 3895-

3962. [https://doi.org/10.1016/S1573-4471\(07\)04061-2](https://doi.org/10.1016/S1573-4471(07)04061-2)

Elen, J., Clarebout, G., Léonard, R., & Lowyck, J. (2007).

Student-centred and teacher-centred learning environments: What students think. *Teaching in Higher Education*, 12(1), 105-117.

<https://doi.org/10.1080/13562510601102339>

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219-245.

<https://doi.org/10.1177/1077800405284363>.CITATIONS

Förster, N., & Souvignier, E. (2014). Learning progress

assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, 91-100.

<https://doi.org/10.1016/j.learninstruc.2014.02.002>

Fowler, S. (2018). Toward a New Curriculum of Leadership

Competencies: Advances in Motivation Science Call for Rethinking Leadership Development. *Advances in Developing Human Resources*. <https://doi.org/10.1177/1523422318756644>

Gaa, J. P. (1973). EFFECTS OF INDIVIDUAL GOAL-SETTING

CONFERENCES ON ACHIEVEMENT , ATTITUDES , AND GOAL-SETTING BEHAVIOR. *The Journal of Experimental Education*, 42(1), 22-28.

Galdas, P. (2017). Revisiting Bias in Qualitative Research:

Reflections on Its Relationship With Funding and Impact.

International Journal of Qualitative Methods, 16(1), 1-2.
<https://doi.org/10.1177/1609406917748992>

Gilley, J. W., Shelton, P. M., & Gilley, A. (2011).
Developmental leadership: A new perspective for human
resource development. *Advances in Developing Human
Resources*, 13(3), 386-405.
<https://doi.org/10.1177/1523422311424264>

Gonçalves, T., Niemivirta, M., & Lemos, M. S. I. of students'
multiple achievement and social goal profiles and
analysis of their stability and adaptability. (2017).
Identification of students' multiple achievement and
social goal profiles and analysis of their stability and
adaptability. *Learning and Individual Differences*, 54,
149-159. <https://doi.org/10.1016/j.lindif.2017.01.019>

Grant, S., Mayo-Wilson, E., Montgomery, P., Macdonald, G.,
Michie, S., Hopewell, S., & Moher, D. (2018). CONSORT-SPI
2018 explanation and elaboration: Guidance for reporting
social and psychological intervention trials. *Trials*, 19,
406. <https://doi.org/10.1186/s13063-018-2735-z>

Grant, S., Montgomery, P., Hopewell, S., Macdonald, G.,
Moher, D., & Mayo-Wilson, E. (2013). Developing a
Reporting Guideline for Social and Psychological
Intervention Trials. *Research on Social Work Practice*,
23(6), 595-602. <https://doi.org/10.1177/1049731513498118>

- Greenier, V. T., & Whitehead, G. E. K. (2016). Towards a Model of Teacher Leadership in ELT: Authentic Leadership in Classroom Practice. *RELC Journal*, 47(1), 79-95.
<https://doi.org/10.1177/0033688216631203>
- Hacking, I. (1988). Telepathy: Origins of Randomization in Experimental Design. *Artifact and Experiment*, 79(3), 427-451. <https://doi.org/10.1007/sl0869-007-9037-x>
- Harrison, H., Birks, M., Franklin, R., & Mills, J. (2017). Case study research: Foundations and methodological orientations. *Forum Qualitative Social Research*, 18(1).
<https://doi.org/10.17169/FQS-18.1.2655>
- Heckman, J. J. (1991). Randomisation and Social Policy Evaluation. *NBER Technical Working Paper Series*.
- Hellen, K., Marcondes, B., Regina, S., & Lucato, R. (2012). Communication and Involvement: What are the Possibilities of Interconnection between Family-school?, 22(51), 91-99.
- Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *American Statistician*, 55(1), 19-24.
<https://doi.org/10.1198/000313001300339897>
- Huh, Y., & Reigeluth, C. M. (2017). Self-Regulated Learning: The Continuous-Change Conceptual Framework and a Vision of New Paradigm, Technology System, and Pedagogical Support. *Journal of Educational Technology Systems*,

46(2), 191-214. <https://doi.org/10.1177/0047239517710769>

Johnson, R. B. B., & Onwuegbuzie, A. J. J. A. J. (2004).

Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26. Retrieved from <http://edr.sagepub.com/content/33/7/14.short>

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, 19(2), 141-184.

<https://doi.org/10.1007/s10648-006-9012-5>

Katyal, K. (2010). Educating teachers in Hong Kong for leadership in the age of the internet: A re-conceptualisation. *Asia-Pacific Journal of Teacher Education*, 38(4), 273-284.

<https://doi.org/10.1080/1359866X.2010.515938>

Katyal, K., & Evers, C. W. (2004). Teacher leadership and autonomous student learning: Adjusting to the new realities. *International Journal of Educational Research*, 41(4-5), 367-382.

<https://doi.org/10.1016/j.ijer.2005.08.007>

Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research*, 86(4), 945-980. <https://doi.org/10.3102/0034654315626800>

Kilday, J. E., Lenser, M. L., & Miller, A. D. (2016). Considering students in teachers' self-efficacy:

Examination of a scale for student-oriented teaching.

Teaching and Teacher Education, 56, 61-71.

<https://doi.org/10.1016/j.tate.2016.01.025>

Komives, S. R., Guthrie, K. L., Hastings, L. J., Hastings, L.

J., Kane, C., & Komives, E. S. R. (2018). Role of

Mentoring , Coaching , and Advising in Developing

Leadership, (158), 23-35.

<https://doi.org/10.1002/yd.20285>

Korthagen, F. A. J. (2010). Situated learning theory and the

pedagogy of teacher education: Towards an integrative

view of teacher behavior and teacher learning. *Teaching*

and Teacher Education, 26(1), 98-106.

<https://doi.org/10.1016/j.tate.2009.05.001>

Korthagen, F. A. J., & Kessels, J. P. A. M. (1999). Linking

Theory and Practice: Changing the Pedagogy of Teacher

Education. *Educational Researcher*, 28(4), 4-17.

<https://doi.org/10.3102/0013189X028004004>

Kraft, M. A., & Blazar, D. (2017). Individualized Coaching to

Improve Teacher Practice Across Grades and Subjects: New

Experimental Evidence. *Educational Policy*, 31(7), 1033-

1068. <https://doi.org/10.1177/0895904816631099>

Lao, A. C. C., Cheng, H. N. H., Huang, M. C. L., Ku, O., &

Chan, T. W. (2017). Examining Motivational Orientation

and Learning Strategies in Computer-Supported Self-

Directed Learning (CS-SDL) for Mathematics: The Perspective of Intrinsic and Extrinsic Goals. *Journal of Educational Computing Research*, 54(8), 1168-1188.

<https://doi.org/10.1177/0735633116651271>

Lee, C. S., Osop, H., Hoe-Lian Goh, D., & Kelni, G. (2017). Making sense of comments on YouTube educational videos: A self-directed learning perspective. *Online Information Review*, 41(5), 611-625. <https://doi.org/10.1108/OIR-09-2016-0274>

Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it. *Educational Technology Research and Development*, 64(4), 707-734.

<https://doi.org/10.1007/s11423-015-9422-5>

Lemos, M. S. De. (1996). Students' and teachers' goals in the classroom. *Learning and Instruction*, 6(2), 151-171.

[https://doi.org/10.1016/0959-4752\(95\)00031-3](https://doi.org/10.1016/0959-4752(95)00031-3)

Lemos, M. S., Gonçalves, T., Lens, W., & Rodrigues, L. P. (2014). The Nature and Dimensions of Achievement Goals: Mastery, Evaluation, Competition, and Self-Presentation Goals. *Spanish Journal of Psychology*, 17(May), E72.

<https://doi.org/10.1017/sjp.2014.77>

Locke, E. a, & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-

- year odyssey. *The American Psychologist*, 57(9), 705-717.
<https://doi.org/10.1037/0003-066X.57.9.705>
- Louzano, P. (2007). *Do Schools Matter in Brazil? Excellence and Equity in Brazilian Primary Education*. Harvard University.
- Marcondes, M. I. (1999). Teacher Education in Brazil. *Journal of Education for Teaching*, 25(3), 203-213.
<https://doi.org/10.1080/02607479919493>
- Mark, M. M. (2015). *Mixed and Multimethods in Predominantly Quantitative Studies, Especially Experiments and Quasi-Experiments*. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*.
<https://doi.org/10.1093/oxfordhb/9780199933624.013.15>
- McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), 210-221.
<https://doi.org/10.1016/j.jdeveco.2012.01.002>
- Moeller, A. J., Theiler, J. M., & Wu, C. (2012). Goal Setting and Student Achievement: A Longitudinal Study. *Modern Language Journal*, 96(2), 153-169.
<https://doi.org/10.1111/j.1540-4781.2011.01231.x>
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research.

International Journal of Qualitative Methods, 1(2), 13-22. <https://doi.org/10.1177/160940690200100202>

Mosse, R., & Sontheimer, L. E. (1996). *Performance Monitoring Indicators Handbook*. World Bank. <https://doi.org/0253-7494>

Mykkänen, A., Perry, N., & Järvelä, S. (2015). Finnish students' reasons for their achievement in classroom activities: focus on features that support self-regulated learning. *Education 3-13*, 4279(May), 1-16. <https://doi.org/10.1080/03004279.2015.1025802>

OECD. (2016). *PISA, Low-Performing Students: Why They Fall Behind and How To Help Them Succeed*. <https://doi.org/http://dx.doi.org/10.1787/9789264250246-en>

Öqvist, A., & Högström, P. (2018). Don ' t Ask Me Why : Preschool Teachers ' Knowledge in Technology as a Determinant of Leadership Behavior. *Journal of Technology Education*, 29(2), 4-19.

Öqvist, A., & Malmström, M. (2016). Teachers' leadership: a maker or a breaker of students' educational motivation. *School Leadership and Management*, 36(4), 365-380. <https://doi.org/10.1080/13632434.2016.1247039>

Öqvist, A., & Malmström, M. (2018). What motivates students? A study on the effects of teacher leadership and

students' self-efficacy. *International Journal of Leadership in Education*, 3124, 1-21.

<https://doi.org/10.1080/13603124.2017.1355480>

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the Science and Practice of Social and Emotional Learning. *Review of Research in Education*, 40(1), 644-681.

<https://doi.org/10.3102/0091732X16673595>

Palesch, Y. Y. (2014). Some common misperceptions about p-values. *Stroke*, 45(12), e244-e246.

<https://doi.org/10.1038/jid.2014.371>

Pintrich, P. R. (1995). Understanding Self-Regulated Learning.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544-555.

<https://doi.org/10.1037/0022-0663.92.3.544>

Radonic, M. (2017). OKR System as the Reference for Personal and Organizational Objectives. *Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal*, 7(2), 28-37.

Rutland, J. H., & Hall, A. (2013). Involving Families in the Assessment Process. *Dialog*, 16(4), 113-120.

Schinkel, W., & Noordegraaf, M. (2011). Professionalism as symbolic capital: Materials for a Bourdieusian theory of

- professionalism. *Comparative Sociology*, 10(1), 67-96.
<https://doi.org/10.1163/156913310X514083>
- Schonert-Reichl, K. A. (2017). Social and Emotional Learning and Teachers. *Social and Emotional Learning and Teachers*, 27(1), 137. Retrieved from www.futureofchildren.org
- Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine*, 1996(14), 727-732.
<https://doi.org/10.7326/0003-4819-152-11-201006010-00232>
- Schunk, D. H. (1990). Goal Setting and Self Efficacy During Self Regulated Learning. *Educational Psychologist*, 25(1), 3-17. <https://doi.org/10.1207/s15326985ep2501>
- Schunk, D. H., & Mullen, C. A. (2013). Toward a Conceptual Model of Mentoring Research: Integration with Self-Regulated Learning. *Educational Psychology Review*, 25(3), 361-389. <https://doi.org/10.1007/s10648-013-9233-3>
- Schwarzenberg, P., Navon, J., Nussbaum, M., Pérez-Sanagustín, M., & Caballero, D. (2018). Learning experience assessment of flipped courses. *Journal of Computing in Higher Education*, 30(2), 237-258.
<https://doi.org/10.1007/s12528-017-9159-8>
- Scott, C., Stone, B., & Dinham, S. (2001). International Patterns of Teacher Discontent, 9(28), 1-18.

- Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation*, 5(9), 11-24. Retrieved from http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/160/186
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *American Education Research Association*, 15(2), 4-14. <https://doi.org/10.1017/CB09781107415324.004>
- Silva, J. L., Bonamino, A. M. C., & Ribeiro, V. M. (2012). Escolas eficazes na educação de jovens e adultos: Estudo de casos na rede municipal do Rio de Janeiro. *Educação Em Revista*, 28, 367-392. <https://doi.org/10.1590/S0102-46982012000200017>
- Souza, P. B. De, & Batista, A. P. (2018). Desempenho Acadêmico e Percepção de Crianças sobre Estilos de Liderança de Professores. *Psicologia Escolar e Educacional*, 22(1), 37-45.
- Stokes, L. (2014). Sample Size Calculation for a Hypothesis Test. *JAMA: The Journal of the American Medical Association*, 1-2. <https://doi.org/10.1001/jama.2014.7195.2>
- Sull, D., & Sull, C. (2018). With Goals, FAST Beats SMART. *MIT Sloan Management Review*, 1-14.

Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., & Furukawa, T. A. (2014). Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Medical Research Methodology*, *14*(1), 1-7.

<https://doi.org/10.1186/1471-2288-14-30>

Thirsk, L. M., & Clark, A. M. (2017). Using Qualitative Research for Complex Interventions: The Contributions of Hermeneutics. *International Journal of Qualitative Methods*, *16*(1), 1-10.

<https://doi.org/10.1177/1609406917721068>

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2008). Achievement goal orientations and subjective well-being: A person-centred analysis. *Learning and Instruction*, *18*(3), 251-266.

<https://doi.org/10.1016/j.learninstruc.2007.05.003>

Usher, E. L., & Pajares, F. (2007). Self-Efficacy for Self-Regulated Learning: A Validation Study. *Educational and Psychological Measurement*, *68*(3), 443-463.

<https://doi.org/10.1177/0013164407308475>

Vedder-Weiss, D., & Fortus, D. (2018). Teachers' Mastery Goals: Using a Self-Report Survey to Study the Relations between Teaching Practices and Students' Motivation for Science Learning. *Research in Science Education*, *48*(1), 181-206. <https://doi.org/10.1007/s11165-016-9565-3>

- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p -Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133.
<https://doi.org/10.1080/00031305.2016.1154108>
- White, H. (2013). An introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness*, 5(1), 30-49.
<https://doi.org/10.1080/19439342.2013.764652>
- Wisdom, J. P., Cavaleri, M. A., Onwuegbuzie, A. J., & Green, C. A. (2012). Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. *Health Services Research*, 47(2), 721-745.
<https://doi.org/10.1111/j.1475-6773.2011.01344.x>
- Wodtke, C. (2015). *Radical Focus: Achieving Your Most Important Goals with Objectives and Key Results*. eBook Kindle.
- Wodtke, C. (2016). *Introduction to OKRs*.
- Yarborough, J. P. (2018). The Role of Coaching in Leadership Development. *New Directions for Student Leadership*, 2018(158), 49-61. <https://doi.org/10.1002/yd.20287>
- Zhou, H., & He, Y. (2018). Comparative Study of OKR and KPI, (Eced), 319-323.
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview, 41(2), 64-70.

- Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology, 39*(2), 145-155.
<https://doi.org/10.1016/j.cedpsych.2014.03.004>
- Zong, X., Zhang, L., & Yao, M. (2017). Parental involvement and Chinese elementary students' achievement goals: the moderating role of parenting style. *Educational Studies, 5698*, 1-16. <https://doi.org/10.1080/03055698.2017.1373634>
- Aghera, A., Emery, M., Bounds, R., Bush, C., Stansfield, B., Gillett, B., & Santen, S. (2018). A Randomized Trial of SMART Goal Enhanced Debriefing after Simulation to Promote Educational Actions. *Western Journal of Emergency Medicine, 19*(January), 112-120.
<https://doi.org/10.5811/westjem.2017.11.36524>
- Altman, D. G. (1985). Comparability of Randomised Groups. *Journal of the Royal Statistical Society. Series D (The Statistician), 34*(1), 125-136.
- Ames, C. (1992). Classrooms: Goals, Structures and Student Motivation. *Journal of Educational Psychology, 84*(3), 261-271.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260-

267. <https://doi.org/10.1037/0022-0663.80.3.260>

- Balart, P., Oosterveen, M., & Webbink, D. (2018). Test scores, noncognitive skills and economic growth. *Economics of Education Review*, 63(October 2017), 134-153. <https://doi.org/10.1016/j.econedurev.2017.12.004>
- Bambirra, R. (2017). Motivation To Learn English As a Foreign Language in Brazil - Giving Voice To a Group of Students At a Public Secondary School. *Linguagem Em Discurso*, 17(2), 215-236. <https://doi.org/10.1590/1982-4017-170204-5316>
- Baxter, P., & Jack, S. (2008). The Qualitative Report Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544-559. <https://doi.org/citeulike-article-id:6670384>
- Berliner, D. C. (2004). Describing the Behavior and Documenting the Accomplishments of Expert Teachers. *Bulletin of Science, Technology & Society*, 24(3), 200-212. <https://doi.org/10.1177/0270467604265535>
- Berson, Y., Halevy, N., Shamir, B., & Erez, M. (2015). Leading from different psychological distances: A construal-level perspective on vision communication , goal setting , and follower motivation. *The Leadership Quarterly*, 26(2), 143-155.

<https://doi.org/10.1016/j.leaqua.2014.07.011>

Brandell, J. R., & Varkas, T. (2001). Narrative Case Studies. In B. A. Thyer (Ed.), *The Handbook of Social Work Research Methods* (pp. 294-307).

British Educational Research Association. (2011). Ethical guidelines for educational research. *British Educational Research Association*, 1-11. <https://doi.org/978-0-946671-32-8>

Bruns, B., Costa, L., & Cunha, N. (2018). Through the looking glass: Can classroom observation and coaching improve teacher performance in Brazil? *Economics of Education Review*, 64(June 2017), 214-250. <https://doi.org/10.1016/j.econedurev.2018.03.003>

Bullock, K., & Wikeley, F. (1999). Improving learning in year 9: Making use of personal learning plans. *Educational Studies*, 25(1), 19-33. <https://doi.org/10.1080/03055699997945>

Cartwright, N. (2011). A philosopher's view of the long road from RCTs to effectiveness. *The Lancet*, 377(9775), 1400-1401. [https://doi.org/10.1016/S0140-6736\(11\)60563-1](https://doi.org/10.1016/S0140-6736(11)60563-1)

Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn't always travel. *Evidence and Policy*, 9(1), 97-112. <https://doi.org/10.1332/174426413X662581>

Cartwright, N., & Hardie, J. (2017). Predicting What Will

- Happen When You Intervene. *Clinical Social Work Journal*, 45(3), 270-279. <https://doi.org/10.1007/s10615-016-0615-0>
- Castro, F. (2018). The OKR Cycle: 3 Steps for OKR Success. Retrieved from <https://felipecastro.com/en/okr-cycle/>
- Childs, A., & McNicholl, J. (2007). Investigating the relationship between subject content knowledge and pedagogical practice through the analysis of classroom discourse. *International Journal of Science Education*, 29(13), 1629-1653. <https://doi.org/10.1080/09500690601180817>
- Cohen, J. (1990a). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312. <https://doi.org/10.1037/0003-066X.45.12.1304>
- Cohen, J. (1990b). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312. <https://doi.org/10.1037/0003-066X.45.12.1304>
- Cohn, L. D., & Decker, B. J. (2003). How Meta-Analysis Increases Statistical Power. *Psychological Methods*, 8(3), 243-253. <https://doi.org/10.1037/1082-989X.8.3.243>
- Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education: Theory and Practice*, 20(1), 73-83. <https://doi.org/10.1007/s10459-014-9509-5>

Day, T., & Tosey, P. (2011). Beyond SMART? A new framework for goal setting. *Curriculum Journal*, 22(4), 515-534.
<https://doi.org/10.1080/09585176.2011.627213>

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine*, 210(October 2017), 2-21.
<https://doi.org/10.1016/j.socscimed.2017.12.005>

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology*, 49(1), 14-23.
<https://doi.org/10.1037/0708-5591.49.1.14>

Deci, E. L., & Ryan, R. M. (2012). *Motivation, Personality, and Development Within Embedded Social Contexts: An Overview of Self-Determination Theory*. *The Oxford Handbook of Human Motivation*.
<https://doi.org/10.1093/oxfordhb/9780195399820.013.0006>

Demidenko, E. (2016). The p-Value You Can't Buy. *American Statistician*, 70(1), 33-38.
<https://doi.org/10.1080/00031305.2015.1069760>

Deole, S. S. (2017). Human Capital Consequences of Violence in Schools: Estimating the Impact of Violence in Schools on Education Outcomes in Brazil. *SSRN Electronic Journal*, 287-310. <https://doi.org/10.2139/ssrn.2925848>

Doerr, J. (2018). *Measure What Matters*. New York:

Portfolio/Penguin.

- Doran, G. T. (1981). There's a S.M.A.R.T. way to write managements's goals and objectives. *Management Review*.
<https://doi.org/10.1177/004057368303900411>
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics*, 4(07), 3895-3962. [https://doi.org/10.1016/S1573-4471\(07\)04061-2](https://doi.org/10.1016/S1573-4471(07)04061-2)
- Elen, J., Clarebout, G., Léonard, R., & Lowyck, J. (2007). Student-centred and teacher-centred learning environments: What students think. *Teaching in Higher Education*, 12(1), 105-117.
<https://doi.org/10.1080/13562510601102339>
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219-245.
<https://doi.org/10.1177/1077800405284363>.CITATIONS
- Förster, N., & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, 91-100.
<https://doi.org/10.1016/j.learninstruc.2014.02.002>
- Fowler, S. (2018). Toward a New Curriculum of Leadership Competencies: Advances in Motivation Science Call for Rethinking Leadership Development. *Advances in Developing*

Human Resources. <https://doi.org/10.1177/1523422318756644>

Gaa, J. P. (1973). EFFECTS OF INDIVIDUAL GOAL-SETTING CONFERENCES ON ACHIEVEMENT , ATTITUDES , AND GOAL-SETTING BEHAVIOR. *The Journal of Experimental Education*, 42(1), 22-28.

Galdas, P. (2017). Revisiting Bias in Qualitative Research: Reflections on Its Relationship With Funding and Impact. *International Journal of Qualitative Methods*, 16(1), 1-2. <https://doi.org/10.1177/1609406917748992>

Gilley, J. W., Shelton, P. M., & Gilley, A. (2011). Developmental leadership: A new perspective for human resource development. *Advances in Developing Human Resources*, 13(3), 386-405. <https://doi.org/10.1177/1523422311424264>

Gonçalves, T., Niemivirta, M., & Lemos, M. S. I. of students' multiple achievement and social goal profiles and analysis of their stability and adaptability. (2017). Identification of students' multiple achievement and social goal profiles and analysis of their stability and adaptability. *Learning and Individual Differences*, 54, 149-159. <https://doi.org/10.1016/j.lindif.2017.01.019>

Grant, S., Mayo-Wilson, E., Montgomery, P., Macdonald, G., Michie, S., Hopewell, S., & Moher, D. (2018). CONSORT-SPI 2018 explanation and elaboration: Guidance for reporting

- social and psychological intervention trials. *Trials*, 19, 406. <https://doi.org/10.1186/s13063-018-2735-z>
- Grant, S., Montgomery, P., Hopewell, S., Macdonald, G., Moher, D., & Mayo-Wilson, E. (2013). Developing a Reporting Guideline for Social and Psychological Intervention Trials. *Research on Social Work Practice*, 23(6), 595-602. <https://doi.org/10.1177/1049731513498118>
- Greenier, V. T., & Whitehead, G. E. K. (2016). Towards a Model of Teacher Leadership in ELT: Authentic Leadership in Classroom Practice. *RELC Journal*, 47(1), 79-95. <https://doi.org/10.1177/0033688216631203>
- Hacking, I. (1988). Telepathy: Origins of Randomization in Experimental Design. *Artifact and Experiment*, 79(3), 427-451. <https://doi.org/10.1007/s10869-007-9037-x>
- Harrison, H., Birks, M., Franklin, R., & Mills, J. (2017). Case study research: Foundations and methodological orientations. *Forum Qualitative Social Research*, 18(1). <https://doi.org/10.17169/FQS-18.1.2655>
- Heckman, J. J. (1991). Randomisation and Social Policy Evaluation. *NBER Technical Working Paper Series*.
- Hellen, K., Marcondes, B., Regina, S., & Lucato, R. (2012). Communication and Involvement: What are the Possibilities of Interconnection between Family-school?, 22(51), 91-99.
- Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power:

The pervasive fallacy of power calculations for data analysis. *American Statistician*, 55(1), 19-24.

<https://doi.org/10.1198/000313001300339897>

Huh, Y., & Reigeluth, C. M. (2017). Self-Regulated Learning: The Continuous-Change Conceptual Framework and a Vision of New Paradigm, Technology System, and Pedagogical Support. *Journal of Educational Technology Systems*, 46(2), 191-214. <https://doi.org/10.1177/00472395177110769>

Johnson, R. B. B., & Onwuegbuzie, A. J. J. A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26. Retrieved from <http://edr.sagepub.com/content/33/7/14.short>

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, 19(2), 141-184.

<https://doi.org/10.1007/s10648-006-9012-5>

Katyal, K. (2010). Educating teachers in Hong Kong for leadership in the age of the internet: A re-conceptualisation. *Asia-Pacific Journal of Teacher Education*, 38(4), 273-284.

<https://doi.org/10.1080/1359866X.2010.515938>

Katyal, K., & Evers, C. W. (2004). Teacher leadership and autonomous student learning: Adjusting to the new realities. *International Journal of Educational Research*,

41(4-5), 367-382.

<https://doi.org/10.1016/j.ijer.2005.08.007>

Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research*, 86(4), 945-980. <https://doi.org/10.3102/0034654315626800>

Kilday, J. E., Lenser, M. L., & Miller, A. D. (2016). Considering students in teachers' self-efficacy: Examination of a scale for student-oriented teaching. *Teaching and Teacher Education*, 56, 61-71. <https://doi.org/10.1016/j.tate.2016.01.025>

Komives, S. R., Guthrie, K. L., Hastings, L. J., Hastings, L. J., Kane, C., & Komives, E. S. R. (2018). Role of Mentoring , Coaching , and Advising in Developing Leadership, (158), 23-35. <https://doi.org/10.1002/yd.20285>

Korthagen, F. A. J. (2010). Situated learning theory and the pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*, 26(1), 98-106. <https://doi.org/10.1016/j.tate.2009.05.001>

Korthagen, F. A. J., & Kessels, J. P. A. M. (1999). Linking Theory and Practice: Changing the Pedagogy of Teacher Education. *Educational Researcher*, 28(4), 4-17. <https://doi.org/10.3102/0013189X028004004>

- Kraft, M. A., & Blazar, D. (2017). Individualized Coaching to Improve Teacher Practice Across Grades and Subjects: New Experimental Evidence. *Educational Policy*, 31(7), 1033-1068. <https://doi.org/10.1177/0895904816631099>
- Lao, A. C. C., Cheng, H. N. H., Huang, M. C. L., Ku, O., & Chan, T. W. (2017). Examining Motivational Orientation and Learning Strategies in Computer-Supported Self-Directed Learning (CS-SDL) for Mathematics: The Perspective of Intrinsic and Extrinsic Goals. *Journal of Educational Computing Research*, 54(8), 1168-1188. <https://doi.org/10.1177/0735633116651271>
- Lee, C. S., Osop, H., Hoe-Lian Goh, D., & Kelni, G. (2017). Making sense of comments on YouTube educational videos: A self-directed learning perspective. *Online Information Review*, 41(5), 611-625. <https://doi.org/10.1108/OIR-09-2016-0274>
- Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it. *Educational Technology Research and Development*, 64(4), 707-734. <https://doi.org/10.1007/s11423-015-9422-5>
- Lemos, M. S. De. (1996). Students' and teachers' goals in the classroom. *Learning and Instruction*, 6(2), 151-171. [https://doi.org/10.1016/0959-4752\(95\)00031-3](https://doi.org/10.1016/0959-4752(95)00031-3)

- Lemos, M. S., Gonçalves, T., Lens, W., & Rodrigues, L. P. (2014). The Nature and Dimensions of Achievement Goals: Mastery, Evaluation, Competition, and Self-Presentation Goals. *Spanish Journal of Psychology*, 17(May), E72. <https://doi.org/10.1017/sjp.2014.77>
- Locke, E. a, & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *The American Psychologist*, 57(9), 705-717. <https://doi.org/10.1037/0003-066X.57.9.705>
- Louzano, P. (2007). *Do Schools Matter in Brazil? Excellence and Equity in Brazilian Primary Education*. Harvard University.
- Marcondes, M. I. (1999). Teacher Education in Brazil. *Journal of Education for Teaching*, 25(3), 203-213. <https://doi.org/10.1080/02607479919493>
- Mark, M. M. (2015). *Mixed and Multimethods in Predominantly Quantitative Studies, Especially Experiments and Quasi-Experiments*. *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*. <https://doi.org/10.1093/oxfordhb/9780199933624.013.15>
- McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), 210-221. <https://doi.org/10.1016/j.jdeveco.2012.01.002>

- Moeller, A. J., Theiler, J. M., & Wu, C. (2012). Goal Setting and Student Achievement: A Longitudinal Study. *Modern Language Journal, 96*(2), 153-169.
<https://doi.org/10.1111/j.1540-4781.2011.01231.x>
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods, 1*(2), 13-22. <https://doi.org/10.1177/160940690200100202>
- Mosse, R., & Sontheimer, L. E. (1996). *Performance Monitoring Indicators Handbook*. World Bank. <https://doi.org/0253-7494>
- Mykkänen, A., Perry, N., & Järvelä, S. (2015). Finnish students' reasons for their achievement in classroom activities: focus on features that support self-regulated learning. *Education 3-13, 4279*(May), 1-16.
<https://doi.org/10.1080/03004279.2015.1025802>
- OECD. (2016). *PISA, Low-Performing Students: Why They Fall Behind and How To Help Them Succeed*.
<https://doi.org/http://dx.doi.org/10.1787/9789264250246-en>
- Öqvist, A., & Högström, P. (2018). Don ' t Ask Me Why: Preschool Teachers ' Knowledge in Technology as a Determinant of Leadership Behavior. *Journal of Technology*

Education, 29(2), 4-19.

Öqvist, A., & Malmström, M. (2016). Teachers' leadership: a maker or a breaker of students' educational motivation. *School Leadership and Management*, 36(4), 365-380.
<https://doi.org/10.1080/13632434.2016.1247039>

Öqvist, A., & Malmström, M. (2018). What motivates students? A study on the effects of teacher leadership and students' self-efficacy. *International Journal of Leadership in Education*, 31(24), 1-21.
<https://doi.org/10.1080/13603124.2017.1355480>

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the Science and Practice of Social and Emotional Learning. *Review of Research in Education*, 40(1), 644-681.
<https://doi.org/10.3102/0091732X16673595>

Palesch, Y. Y. (2014). Some common misperceptions about p-values. *Stroke*, 45(12), e244-e246.
<https://doi.org/10.1038/jid.2014.371>

Pintrich, P. R. (1995). Understanding Self-Regulated Learning.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544-555.
<https://doi.org/10.1037/0022-0663.92.3.544>

- Radonic, M. (2017). OKR System as the Reference for Personal and Organizational Objectives. *Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal*, 7(2), 28-37.
- Rutland, J. H., & Hall, A. (2013). Involving Families in the Assessment Process. *Dialog*, 16(4), 113-120.
- Schinkel, W., & Noordegraaf, M. (2011). Professionalism as symbolic capital: Materials for a Bourdieusian theory of professionalism. *Comparative Sociology*, 10(1), 67-96.
<https://doi.org/10.1163/156913310X514083>
- Schonert-Reichl, K. A. (2017). Social and Emotional Learning and Teachers. *Social and Emotional Learning and Teachers*, 27(1), 137. Retrieved from www.futureofchildren.org
- Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine*, 152(14), 727-732.
<https://doi.org/10.7326/0003-4819-152-11-201006010-00232>
- Schunk, D. H. (1990). Goal Setting and Self Efficacy During Self Regulated Learning. *Educational Psychologist*, 25(1), 3-17. <https://doi.org/10.1207/s15326985ep2501>
- Schunk, D. H., & Mullen, C. A. (2013). Toward a Conceptual Model of Mentoring Research: Integration with Self-Regulated Learning. *Educational Psychology Review*, 25(3), 361-389. <https://doi.org/10.1007/s10648-013-9233-3>

- Schwarzenberg, P., Navon, J., Nussbaum, M., Pérez-Sanagustín, M., & Caballero, D. (2018). Learning experience assessment of flipped courses. *Journal of Computing in Higher Education, 30*(2), 237-258.
<https://doi.org/10.1007/s12528-017-9159-8>
- Scott, C., Stone, B., & Dinham, S. (2001). International Patterns of Teacher Discontent, *9*(28), 1-18.
- Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation, 5*(9), 11-24. Retrieved from
http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/160/186
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *American Education Research Association, 15*(2), 4-14. <https://doi.org/10.1017/CBO9781107415324.004>
- Silva, J. L., Bonamino, A. M. C., & Ribeiro, V. M. (2012). Escolas eficazes na educação de jovens e adultos: Estudo de casos na rede municipal do Rio de Janeiro. *Educação Em Revista, 28*, 367-392. <https://doi.org/10.1590/S0102-46982012000200017>
- Souza, P. B. De, & Batista, A. P. (2018). Desempenho Acadêmico e Percepção de Crianças sobre Estilos de Liderança de Professores. *Psicologia Escolar e*

Educacional, 22(1), 37-45.

Stokes, L. (2014). Sample Size Calculation for a Hypothesis Test. *JAMA: The Journal of the American Medical Association*, 1-2.

<https://doi.org/10.1001/jama.2014.7195.2>

Sull, D., & Sull, C. (2018). With Goals, FAST Beats SMART. *MIT Sloan Management Review*, 1-14.

Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., & Furukawa, T. A. (2014). Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Medical Research Methodology*, 14(1), 1-7.

<https://doi.org/10.1186/1471-2288-14-30>

Thirsk, L. M., & Clark, A. M. (2017). Using Qualitative Research for Complex Interventions: The Contributions of Hermeneutics. *International Journal of Qualitative Methods*, 16(1), 1-10.

<https://doi.org/10.1177/1609406917721068>

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2008). Achievement goal orientations and subjective well-being: A person-centred analysis. *Learning and Instruction*, 18(3), 251-266.

<https://doi.org/10.1016/j.learninstruc.2007.05.003>

Usher, E. L., & Pajares, F. (2007). Self-Efficacy for Self-

Regulated Learning: A Validation Study. *Educational and Psychological Measurement*, 68(3), 443-463.

<https://doi.org/10.1177/0013164407308475>

Vedder-Weiss, D., & Fortus, D. (2018). Teachers' Mastery Goals: Using a Self-Report Survey to Study the Relations between Teaching Practices and Students' Motivation for Science Learning. *Research in Science Education*, 48(1), 181-206. <https://doi.org/10.1007/s11165-016-9565-3>

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on *p*-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133.

<https://doi.org/10.1080/00031305.2016.1154108>

White, H. (2013). An introduction to the use of randomised control trials to evaluate development interventions. *Journal of Development Effectiveness*, 5(1), 30-49.

<https://doi.org/10.1080/19439342.2013.764652>

Wisdom, J. P., Cavaleri, M. A., Onwuegbuzie, A. J., & Green, C. A. (2012). Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. *Health Services Research*, 47(2), 721-745.

<https://doi.org/10.1111/j.1475-6773.2011.01344.x>

Wodtke, C. (2015). *Radical Focus: Achieving Your Most Important Goals with Objectives and Key Results*. eBook Kindle.

- Wodtke, C. (2016). *Introduction to OKRs*.
- Yarborough, J. P. (2018). The Role of Coaching in Leadership Development. *New Directions for Student Leadership*, 2018(158), 49-61. <https://doi.org/10.1002/yd.20287>
- Zhou, H., & He, Y. (2018). Comparative Study of OKR and KPI, (Eced), 319-323.
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner : An Overview, 41(2), 64-70.
- Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*, 39(2), 145-155. <https://doi.org/10.1016/j.cedpsych.2014.03.004>
- Zong, X., Zhang, L., & Yao, M. (2017). Parental involvement and Chinese elementary students' achievement goals: the moderating role of parenting style. *Educational Studies*, 5698, 1-16. <https://doi.org/10.1080/03055698.2017.1373634>
- Aghera, A., Emery, M., Bounds, R., Bush, C., Stansfield, B., Gillett, B., & Santen, S. (2018). A Randomized Trial of SMART Goal Enhanced Debriefing after Simulation to Promote Educational Actions. *Western Journal of Emergency Medicine*, 19(January), 112-120. <https://doi.org/10.5811/westjem.2017.11.36524>
- Altman, D. G. (1985). Comparability of Randomised Groups.

Journal of the Royal Statistical Society. Series D (The Statistician), 34(1), 125-136.

Ames, C. (1992). Classrooms: Goals, Structures and Student Motivation. *Journal of Educational Psychology*, 84(3), 261-271.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80(3), 260-267. <https://doi.org/10.1037/0022-0663.80.3.260>

Balart, P., Oosterveen, M., & Webbink, D. (2018). Test scores, noncognitive skills and economic growth. *Economics of Education Review*, 63(October 2017), 134-153. <https://doi.org/10.1016/j.econedurev.2017.12.004>

Bambirra, R. (2017). Motivation To Learn English As a Foreign Language in Brazil - Giving Voice To a Group of Students At a Public Secondary School. *Linguagem Em Discurso*, 17(2), 215-236. <https://doi.org/10.1590/1982-4017-170204-5316>

Baxter, P., & Jack, S. (2008). The Qualitative Report Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544-559. <https://doi.org/citeulike-article-id:6670384>

Berliner, D. C. (2004). Describing the Behavior and

Documenting the Accomplishments of Expert Teachers.
Bulletin of Science, Technology & Society, 24(3), 200-
212. <https://doi.org/10.1177/0270467604265535>

Berson, Y., Halevy, N., Shamir, B., & Erez, M. (2015).
Leading from different psychological distances: A
construal-level perspective on vision communication,
goal setting, and follower motivation. *The Leadership
Quarterly*, 26(2), 143-155.
<https://doi.org/10.1016/j.leaqua.2014.07.011>

Brandell, J. R., & Varkas, T. (2001). Narrative Case Studies.
In B. A. Thyer (Ed.), *The Handbook of Social Work
Research Methods* (pp. 294-307).

British Educational Research Association. (2011). Ethical
guidelines for educational research. *British Educational
Research Association*, 1-11. [https://doi.org/978-0-946671-
32-8](https://doi.org/978-0-946671-32-8)

Bruns, B., Costa, L., & Cunha, N. (2018). Through the looking
glass: Can classroom observation and coaching improve
teacher performance in Brazil? *Economics of Education
Review*, 64(June 2017), 214-250.
<https://doi.org/10.1016/j.econedurev.2018.03.003>

Bullock, K., & Wikeley, F. (1999). Improving learning in year
9: Making use of personal learning plans. *Educational
Studies*, 25(1), 19-33.

<https://doi.org/10.1080/03055699997945>

- Cartwright, N. (2011). A philosopher's view of the long road from RCTs to effectiveness. *The Lancet*, 377(9775), 1400-1401. [https://doi.org/10.1016/S0140-6736\(11\)60563-1](https://doi.org/10.1016/S0140-6736(11)60563-1)
- Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn't always travel. *Evidence and Policy*, 9(1), 97-112. <https://doi.org/10.1332/174426413X662581>
- Cartwright, N., & Hardie, J. (2017). Predicting What Will Happen When You Intervene. *Clinical Social Work Journal*, 45(3), 270-279. <https://doi.org/10.1007/s10615-016-0615-0>
- Castro, F. (2018). The OKR Cycle: 3 Steps for OKR Success. Retrieved from <https://felipecastro.com/en/okr-cycle/>
- Childs, A., & McNicholl, J. (2007). Investigating the relationship between subject content knowledge and pedagogical practice through the analysis of classroom discourse. *International Journal of Science Education*, 29(13), 1629-1653.
<https://doi.org/10.1080/09500690601180817>
- Cohen, J. (1990a). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312.
<https://doi.org/10.1037/0003-066X.45.12.1304>
- Cohen, J. (1990b). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312.
<https://doi.org/10.1037/0003-066X.45.12.1304>

- Cohn, L. D., & Decker, B. J. (2003). How Meta-Analysis Increases Statistical Power. *Psychological Methods, 8*(3), 243-253. <https://doi.org/10.1037/1082-989X.8.3.243>
- Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education: Theory and Practice, 20*(1), 73-83. <https://doi.org/10.1007/s10459-014-9509-5>
- Day, T., & Tosey, P. (2011). Beyond SMART? A new framework for goal setting. *Curriculum Journal, 22*(4), 515-534. <https://doi.org/10.1080/09585176.2011.627213>
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine, 210*(October 2017), 2-21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology, 49*(1), 14-23. <https://doi.org/10.1037/0708-5591.49.1.14>
- Deci, E. L., & Ryan, R. M. (2012). *Motivation, Personality, and Development Within Embedded Social Contexts: An Overview of Self-Determination Theory. The Oxford Handbook of Human Motivation.* <https://doi.org/10.1093/oxfordhb/9780195399820.013.0006>

- Demidenko, E. (2016). The p-Value You Can't Buy. *American Statistician*, 70(1), 33-38.
<https://doi.org/10.1080/00031305.2015.1069760>
- Deole, S. S. (2017). Human Capital Consequences of Violence in Schools: Estimating the Impact of Violence in Schools on Education Outcomes in Brazil. *SSRN Electronic Journal*, 287-310. <https://doi.org/10.2139/ssrn.2925848>
- Doerr, J. (2018). *Measure What Matters*. New York: Portfolio/Penguin.
- Doran, G. T. (1981). There's a S.M.A.R.T. way to write managements's goals and objectives. *Management Review*.
<https://doi.org/10.1177/004057368303900411>
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics*, 4(07), 3895-3962. [https://doi.org/10.1016/S1573-4471\(07\)04061-2](https://doi.org/10.1016/S1573-4471(07)04061-2)
- Elen, J., Clarebout, G., Léonard, R., & Lowyck, J. (2007). Student-centred and teacher-centred learning environments: What students think. *Teaching in Higher Education*, 12(1), 105-117.
<https://doi.org/10.1080/13562510601102339>
- Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219-245.
<https://doi.org/10.1177/1077800405284363>.CITATIONS

- Förster, N., & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction, 32*, 91-100.
<https://doi.org/10.1016/j.learninstruc.2014.02.002>
- Fowler, S. (2018). Toward a New Curriculum of Leadership Competencies: Advances in Motivation Science Call for Rethinking Leadership Development. *Advances in Developing Human Resources*. <https://doi.org/10.1177/1523422318756644>
- Gaa, J. P. (1973). EFFECTS OF INDIVIDUAL GOAL-SETTING CONFERENCES ON ACHIEVEMENT , ATTITUDES , AND GOAL-SETTING BEHAVIOR. *The Journal of Experimental Education, 42*(1), 22-28.
- Galdas, P. (2017). Revisiting Bias in Qualitative Research: Reflections on Its Relationship With Funding and Impact. *International Journal of Qualitative Methods, 16*(1), 1-2.
<https://doi.org/10.1177/1609406917748992>
- Gilley, J. W., Shelton, P. M., & Gilley, A. (2011). Developmental leadership: A new perspective for human resource development. *Advances in Developing Human Resources, 13*(3), 386-405.
<https://doi.org/10.1177/1523422311424264>
- Gonçalves, T., Niemivirta, M., & Lemos, M. S. I. of students' multiple achievement and social goal profiles and

- analysis of their stability and adaptability. (2017). Identification of students' multiple achievement and social goal profiles and analysis of their stability and adaptability. *Learning and Individual Differences*, 54, 149-159. <https://doi.org/10.1016/j.lindif.2017.01.019>
- Grant, S., Mayo-Wilson, E., Montgomery, P., Macdonald, G., Michie, S., Hopewell, S., & Moher, D. (2018). CONSORT-SPI 2018 explanation and elaboration: Guidance for reporting social and psychological intervention trials. *Trials*, 19, 406. <https://doi.org/10.1186/s13063-018-2735-z>
- Grant, S., Montgomery, P., Hopewell, S., Macdonald, G., Moher, D., & Mayo-Wilson, E. (2013). Developing a Reporting Guideline for Social and Psychological Intervention Trials. *Research on Social Work Practice*, 23(6), 595-602. <https://doi.org/10.1177/1049731513498118>
- Greenier, V. T., & Whitehead, G. E. K. (2016). Towards a Model of Teacher Leadership in ELT: Authentic Leadership in Classroom Practice. *RELC Journal*, 47(1), 79-95. <https://doi.org/10.1177/0033688216631203>
- Hacking, I. (1988). Telepathy: Origins of Randomization in Experimental Design. *Artifact and Experiment*, 79(3), 427-451. <https://doi.org/10.1007/s10869-007-9037-x>
- Harrison, H., Birks, M., Franklin, R., & Mills, J. (2017). Case study research: Foundations and methodological

- orientations. *Forum Qualitative Social Research*, 18(1).
<https://doi.org/10.17169/FQS-18.1.2655>
- Heckman, J. J. (1991). Randomisation and Social Policy Evaluation. *NBER Technical Working Paper Series*.
- Hellen, K., Marcondes, B., Regina, S., & Lucato, R. (2012). Communication and Involvement: What are the Possibilities of Interconnection between Family-school?, 22(51), 91-99.
- Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *American Statistician*, 55(1), 19-24.
<https://doi.org/10.1198/000313001300339897>
- Huh, Y., & Reigeluth, C. M. (2017). Self-Regulated Learning: The Continuous-Change Conceptual Framework and a Vision of New Paradigm, Technology System, and Pedagogical Support. *Journal of Educational Technology Systems*, 46(2), 191-214. <https://doi.org/10.1177/0047239517710769>
- Johnson, R. B. B., & Onwuegbuzie, A. J. J. A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26. Retrieved from <http://edr.sagepub.com/content/33/7/14.short>
- Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, 19(2), 141-184.
<https://doi.org/10.1007/s10648-006-9012-5>

Katyal, K. (2010). Educating teachers in Hong Kong for leadership in the age of the internet: A re-conceptualisation. *Asia-Pacific Journal of Teacher Education*, 38(4), 273-284.

<https://doi.org/10.1080/1359866X.2010.515938>

Katyal, K., & Evers, C. W. (2004). Teacher leadership and autonomous student learning: Adjusting to the new realities. *International Journal of Educational Research*, 41(4-5), 367-382.

<https://doi.org/10.1016/j.ijer.2005.08.007>

Kennedy, M. M. (2016). How Does Professional Development Improve Teaching? *Review of Educational Research*, 86(4), 945-980. <https://doi.org/10.3102/0034654315626800>

Kilday, J. E., Lenser, M. L., & Miller, A. D. (2016). Considering students in teachers' self-efficacy: Examination of a scale for student-oriented teaching. *Teaching and Teacher Education*, 56, 61-71.

<https://doi.org/10.1016/j.tate.2016.01.025>

Komives, S. R., Guthrie, K. L., Hastings, L. J., Hastings, L. J., Kane, C., & Komives, E. S. R. (2018). Role of Mentoring , Coaching , and Advising in Developing Leadership, (158), 23-35.

<https://doi.org/10.1002/yd.20285>

Korthagen, F. A. J. (2010). Situated learning theory and the

pedagogy of teacher education: Towards an integrative view of teacher behavior and teacher learning. *Teaching and Teacher Education*, 26(1), 98-106.

<https://doi.org/10.1016/j.tate.2009.05.001>

Korthagen, F. A. J., & Kessels, J. P. A. M. (1999). Linking Theory and Practice: Changing the Pedagogy of Teacher Education. *Educational Researcher*, 28(4), 4-17.

<https://doi.org/10.3102/0013189X028004004>

Kraft, M. A., & Blazar, D. (2017). Individualized Coaching to Improve Teacher Practice Across Grades and Subjects: New Experimental Evidence. *Educational Policy*, 31(7), 1033-1068. <https://doi.org/10.1177/0895904816631099>

Lao, A. C. C., Cheng, H. N. H., Huang, M. C. L., Ku, O., & Chan, T. W. (2017). Examining Motivational Orientation and Learning Strategies in Computer-Supported Self-Directed Learning (CS-SDL) for Mathematics: The Perspective of Intrinsic and Extrinsic Goals. *Journal of Educational Computing Research*, 54(8), 1168-1188.

<https://doi.org/10.1177/0735633116651271>

Lee, C. S., Osop, H., Hoe-Lian Goh, D., & Kelni, G. (2017). Making sense of comments on YouTube educational videos: A self-directed learning perspective. *Online Information Review*, 41(5), 611-625. <https://doi.org/10.1108/OIR-09-2016-0274>

- Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: own it, learn it, and share it. *Educational Technology Research and Development*, 64(4), 707-734.
<https://doi.org/10.1007/s11423-015-9422-5>
- Lemos, M. S. De. (1996). Students' and teachers' goals in the classroom. *Learning and Instruction*, 6(2), 151-171.
[https://doi.org/10.1016/0959-4752\(95\)00031-3](https://doi.org/10.1016/0959-4752(95)00031-3)
- Lemos, M. S., Gonçalves, T., Lens, W., & Rodrigues, L. P. (2014). The Nature and Dimensions of Achievement Goals: Mastery, Evaluation, Competition, and Self-Presentation Goals. *Spanish Journal of Psychology*, 17(May), E72.
<https://doi.org/10.1017/sjp.2014.77>
- Locke, E. a, & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *The American Psychologist*, 57(9), 705-717.
<https://doi.org/10.1037/0003-066X.57.9.705>
- Louzano, P. (2007). *Do Schools Matter in Brazil? Excellence and Equity in Brazilian Primary Education*. Harvard University.
- Marcondes, M. I. (1999). Teacher Education in Brazil. *Journal of Education for Teaching*, 25(3), 203-213.
<https://doi.org/10.1080/02607479919493>
- Mark, M. M. (2015). *Mixed and Multimethods in Predominantly*

Quantitative Studies, Especially Experiments and Quasi-Experiments. The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry.

<https://doi.org/10.1093/oxfordhb/9780199933624.013.15>

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), 210-221.

<https://doi.org/10.1016/j.jdeveco.2012.01.002>

Moeller, A. J., Theiler, J. M., & Wu, C. (2012). Goal Setting and Student Achievement: A Longitudinal Study. *Modern Language Journal*, 96(2), 153-169.

<https://doi.org/10.1111/j.1540-4781.2011.01231.x>

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods*, 1(2), 13-22. <https://doi.org/10.1177/160940690200100202>

Mosse, R., & Sontheimer, L. E. (1996). *Performance Monitoring Indicators Handbook*. World Bank. <https://doi.org/0253-7494>

Mykkänen, A., Perry, N., & Järvelä, S. (2015). Finnish students' reasons for their achievement in classroom activities: focus on features that support self-regulated learning. *Education 3-13*, 4279(May), 1-16.

<https://doi.org/10.1080/03004279.2015.1025802>

OECD. (2016). *PISA, Low-Performing Students: Why They Fall Behind and How To Help Them Succeed*.

<https://doi.org/http://dx.doi.org/10.1787/9789264250246-en>

Öqvist, A., & Högström, P. (2018). Don ' t Ask Me Why :
Preschool Teachers ' Knowledge in Technology as a
Determinant of Leadership Behavior. *Journal of Technology
Education, 29*(2), 4-19.

Öqvist, A., & Malmström, M. (2016). Teachers' leadership: a
maker or a breaker of students' educational motivation.
School Leadership and Management, 36(4), 365-380.

<https://doi.org/10.1080/13632434.2016.1247039>

Öqvist, A., & Malmström, M. (2018). What motivates students?
A study on the effects of teacher leadership and
students' self-efficacy. *International Journal of
Leadership in Education, 31*24, 1-21.

<https://doi.org/10.1080/13603124.2017.1355480>

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S.,
& Weissberg, R. P. (2016). Advancing the Science and
Practice of Social and Emotional Learning. *Review of
Research in Education, 40*(1), 644-681.

<https://doi.org/10.3102/0091732X16673595>

Palesch, Y. Y. (2014). Some common misperceptions about p-

values. *Stroke*, 45(12), e244-e246.

<https://doi.org/10.1038/jid.2014.371>

Pintrich, P. R. (1995). Understanding Self-Regulated Learning.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544-555.

<https://doi.org/10.1037/0022-0663.92.3.544>

Radonic, M. (2017). OKR System as the Reference for Personal and Organizational Objectives. *Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal*, 7(2), 28-37.

Rutland, J. H., & Hall, A. (2013). Involving Families in the Assessment Process. *Dialog*, 16(4), 113-120.

Schinkel, W., & Noordegraaf, M. (2011). Professionalism as symbolic capital: Materials for a Bourdieusian theory of professionalism. *Comparative Sociology*, 10(1), 67-96.

<https://doi.org/10.1163/156913310X514083>

Schonert-Reichl, K. A. (2017). Social and Emotional Learning and Teachers. *Social and Emotional Learning and Teachers*, 27(1), 137. Retrieved from www.futureofchildren.org

Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine*, 1996(14), 727-732.

<https://doi.org/10.7326/0003-4819-152-11-201006010-00232>

- Schunk, D. H. (1990). Goal Setting and Self Efficacy During Self Regulated Learning. *Educational Psychologist*, 25(1), 3-17. <https://doi.org/10.1207/s15326985ep2501>
- Schunk, D. H., & Mullen, C. A. (2013). Toward a Conceptual Model of Mentoring Research: Integration with Self-Regulated Learning. *Educational Psychology Review*, 25(3), 361-389. <https://doi.org/10.1007/s10648-013-9233-3>
- Schwarzenberg, P., Navon, J., Nussbaum, M., Pérez-Sanagustín, M., & Caballero, D. (2018). Learning experience assessment of flipped courses. *Journal of Computing in Higher Education*, 30(2), 237-258. <https://doi.org/10.1007/s12528-017-9159-8>
- Scott, C., Stone, B., & Dinham, S. (2001). International Patterns of Teacher Discontent, 9(28), 1-18.
- Scriven, M. (2008). A summative evaluation of RCT methodology: and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation*, 5(9), 11-24. Retrieved from http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/160/186
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *American Education Research Association*, 15(2), 4-14. <https://doi.org/10.1017/CBO9781107415324.004>

- Silva, J. L., Bonamino, A. M. C., & Ribeiro, V. M. (2012). Escolas eficazes na educação de jovens e adultos: Estudo de casos na rede municipal do Rio de Janeiro. *Educação Em Revista, 28*, 367-392. <https://doi.org/10.1590/S0102-46982012000200017>
- Souza, P. B. De, & Batista, A. P. (2018). Desempenho Acadêmico e Percepção de Crianças sobre Estilos de Liderança de Professores. *Psicologia Escolar e Educacional, 22*(1), 37-45.
- Stokes, L. (2014). Sample Size Calculation for a Hypothesis Test. *JAMA: The Journal of the American Medical Association, 1-2*.
<https://doi.org/10.1001/jama.2014.7195.2>
- Sull, D., & Sull, C. (2018). With Goals, FAST Beats SMART. *MIT Sloan Management Review, 1-14*.
- Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., & Furukawa, T. A. (2014). Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Medical Research Methodology, 14*(1), 1-7.
<https://doi.org/10.1186/1471-2288-14-30>
- Thirsk, L. M., & Clark, A. M. (2017). Using Qualitative Research for Complex Interventions: The Contributions of Hermeneutics. *International Journal of Qualitative*

Methods, 16(1), 1-10.

<https://doi.org/10.1177/1609406917721068>

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2008).

Achievement goal orientations and subjective well-being:

A person-centred analysis. *Learning and Instruction*,

18(3), 251-266.

<https://doi.org/10.1016/j.learninstruc.2007.05.003>

Usher, E. L., & Pajares, F. (2007). Self-Efficacy for Self-

Regulated Learning: A Validation Study. *Educational and*

Psychological Measurement, 68(3), 443-463.

<https://doi.org/10.1177/0013164407308475>

Vedder-Weiss, D., & Fortus, D. (2018). Teachers' Mastery

Goals: Using a Self-Report Survey to Study the Relations

between Teaching Practices and Students' Motivation for

Science Learning. *Research in Science Education*, 48(1),

181-206. <https://doi.org/10.1007/s11165-016-9565-3>

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's

Statement on p -Values: Context, Process, and Purpose.

The American Statistician, 70(2), 129-133.

<https://doi.org/10.1080/00031305.2016.1154108>

White, H. (2013). An introduction to the use of randomised

control trials to evaluate development interventions.

Journal of Development Effectiveness, 5(1), 30-49.

<https://doi.org/10.1080/19439342.2013.764652>


- Wisdom, J. P., Cavaleri, M. A., Onwuegbuzie, A. J., & Green, C. A. (2012). Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. *Health Services Research, 47*(2), 721-745. <https://doi.org/10.1111/j.1475-6773.2011.01344.x>
- Wodtke, C. (2015). *Radical Focus: Achieving Your Most Important Goals with Objectives and Key Results*. eBook Kindle.
- Wodtke, C. (2016). *Introduction to OKRs*.
- Yarborough, J. P. (2018). The Role of Coaching in Leadership Development. *New Directions for Student Leadership, 2018*(158), 49-61. <https://doi.org/10.1002/yd.20287>
- Zhou, H., & He, Y. (2018). Comparative Study of OKR and KPI, (Eced), 319-323.
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview, *41*(2), 64-70.
- Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology, 39*(2), 145-155. <https://doi.org/10.1016/j.cedpsych.2014.03.004>
- Zong, X., Zhang, L., & Yao, M. (2017). Parental involvement and Chinese elementary students' achievement goals: the moderating role of parenting style. *Educational Studies,*

Appendix A

Sequence of slides used in the lecture on OKR.

1 MARINA PEDRO LAURA GUILHERME	2 Aprendendo a Programar <small>Aula 0</small>	3 objetivo	4 objetivo escrever um software
5 objetivo escrever um software (sem usar um computador)	6 ?	7 ? um software é um texto	8 ? um software é um texto escrito em uma linguagem
9 ? um software é um texto escrito em uma linguagem que um computador sabe ler	10 ? um software é um texto	11 ? um software é um texto que descreve uma tarefa	12 objetivo escrever um software

13 objetivo	14 objetivo descrever uma tarefa	15 ?	16 pergunta
17 pergunta como vamos saber se atingimos o objetivo?	18 ?	19 ? resultado-chave	20 ? um software é um texto que descreve uma tarefa
21 ? resultado-chave	22 ? resultado-chave (pausa pra reflexao)	23 tarefa procurar uma palavra no dicionario	24 exemplo do professor

<p>Passo 1: Abre dicionario pagina 1 Passo 2: Procura palavra na pagina Passo 3: Se viu a palavra LE_SIGNIFICADO_FIM Passo 4: Se ultima pagina NAO_ACHOU_FIM Passo 5: Vira a pagina e vai para Passo 2</p>	<p> tarefa procurar uma palavra no dicionario</p>	<p> tarefa procurar uma palavra no dicionario usando OUTRO método</p>	<p> ? (15 min)</p>
25	26	27	28
<p>professor o resultado vale nota?</p>	<p>professor o resultado vale nota? não. o resultado não.</p>	<p>professor o resultado vale nota? não. o resultado não. querer alcançar vale 20%</p>	<p>OKR Objectives and Key Results</p>
29	30	31	32
<p>Qual a importância de traçar objetivos?</p> <ul style="list-style-type: none"> • Um dos processos fundamentais da auto-regulação do aprendizado • Auto-regulação tem correlação positiva com resultado acadêmico • Um dos caminhos mais pavimentados para crescimento de indicadores tipo IDEB e PISA é o aumento da auto-regulação dos alunos • A liderança do professor é crucial. A explicitação nem tanto. 	<p>O que é OKR?</p> <ul style="list-style-type: none"> • Método para gerenciamento de objetivos que foi popularizado pelo Google. • Simples. Ciclo Rápido. Transparente. Qualitativo e Quantitativo. • Objetivos tem de ser desacoplados de remuneração. No nosso caso eles tem de ser desacoplados de nota. • Podemos ter várias cadências. OKRs para um tópico como eu demonstrei hoje ou para um período maior. 	<p>E-OKR</p> <ul style="list-style-type: none"> • OKR para Educação • Meu projeto de pesquisa no Departamento de Educação da Universidade de Colômbia • Pegar a BNCC (BENEFÍCIOS) e criar para cada tópico um conjunto de sugestões de objetivos e resultados-chave que o professor poderá propor para os alunos 	
33	34	35	36

Appendix B

Filled RCT questionnaires - Pre (1) and Post-Lecture (2)
(sample)

Questionário 1

1

1	Eu tenho <u>40</u> % de certeza que eu consigo ajudar o aluno a identificar as melhores estratégias de aprendizagem
2	Eu tenho <u>70</u> % de certeza de que eu consigo ajudar o aluno a utilizar estratégias de aprendizagem que serão úteis para ele nos outros anos
3	Eu tenho <u>60</u> % de certeza de que eu consigo criar oportunidades para que o aluno estabeleça seus próprios objetivos
4	Eu tenho <u>80</u> % de certeza de que eu consigo fazer o aluno compreender que usar corretamente o material didático é importante para que ele atinja os objetivos
5	Eu tenho <u>70</u> % de certeza de que eu consigo explicar para o aluno o motivo de cada atividade ser relevante
6	Eu tenho <u>50</u> % de certeza de que eu consigo criar oportunidades para que o aluno faça escolhas relacionadas com o próprio aprendizado
7	Eu tenho <u>40</u> % de certeza de que eu consigo ajudar o aluno a entender como ele está se saindo na minha matéria
8	Eu tenho <u>60</u> % de certeza de que eu consigo ajudar o aluno que tem mais dificuldade a identificar estratégias de aprendizagem que o auxiliem na superação dos obstáculos
9	No que se refere a conseguir fazer o aluno relacionar o conteúdo com outras matérias e assuntos, eu sou nota <u>9</u> (1-10)
10	No que se refere a conseguir criar muitas oportunidades para que o aluno perceba a importância daquele conhecimento, eu sou nota <u>8</u> (1-10)
11	No que se refere a conseguir levar em consideração, sempre que explico algo, o que o aluno já sabe (ou acha que sabe) sobre o assunto, eu sou nota <u>7</u> (1-10)

Questionário 2



1	Eu tenho <u>60</u> % de certeza de que consigo levar em consideração, sempre que explico algo, aquilo que o aluno já sabe (ou acha que sabe) sobre o assunto
2	Quanto a conseguir ajudar o aluno a adotar estratégias de aprendizagem que serão úteis nos próximos anos, eu sou nota <u>7</u> (1-10)
3	No que se refere a conseguir ajudar o aluno a identificar as melhores estratégias de aprendizagem, eu sou nota <u>7</u> (1-10)
4	Quanto a conseguir fazer o aluno entender como ele esta se saindo na minha matéria, eu sou nota <u>8</u> (1-10)
5	Quanto a conseguir fazer o aluno compreender que o uso do material didático é importante para que ele atinja os objetivos, eu sou nota <u>8</u> (1-10)
6	No que se refere a conseguir ajudar o aluno que tem mais dificuldade a identificar estratégias de aprendizagem que o auxiliem na superação dos obstáculos, eu sou nota <u>7</u> (1-10)
7	No que se refere a conseguir criar oportunidades para que o aluno faça escolhas referentes ao próprio aprendizado, eu sou nota <u>6</u> (1-10)
8	Quanto a conseguir criar oportunidades para que o aluno estabeleça seus próprios objetivos de aprendizagem, eu sou nota <u>7</u> (1-10)
9	Eu tenho <u>100</u> % de certeza que consigo fazer o aluno relacionar o conteúdo com outras matérias e assuntos
10	Eu tenho <u>70</u> % de certeza de que eu consigo criar muitas oportunidades para que o aluno perceba que aquele conhecimento é relevante
11	Quanto a conseguir explicar para o aluno o motivo de cada atividade ser relevante, eu sou nota <u>8</u> (1-10)

Questionário 1

6

1	Eu tenho <u>45</u> % de certeza que eu consigo ajudar o aluno a identificar as melhores estratégias de aprendizagem
2	Eu tenho <u>90</u> % de certeza de que eu consigo ajudar o aluno a utilizar estratégias de aprendizagem que serão úteis para ele nos outros anos
3	Eu tenho <u>96</u> % de certeza de que eu consigo criar oportunidades para que o aluno estabeleça seus próprios objetivos
4	Eu tenho <u>96</u> % de certeza de que eu consigo fazer o aluno compreender que usar corretamente o material didático é importante para que ele atinja os objetivos
5	Eu tenho <u>98</u> % de certeza de que eu consigo explicar para o aluno o motivo de cada atividade ser relevante
6	Eu tenho <u>95</u> % de certeza de que eu consigo criar oportunidades para que o aluno faça escolhas relacionadas com o próprio aprendizado
7	Eu tenho <u>98</u> % de certeza de que eu consigo ajudar o aluno a entender como ele está se saindo na minha matéria
8	Eu tenho <u>92</u> % de certeza de que eu consigo ajudar o aluno que tem mais dificuldade a identificar estratégias de aprendizagem que o auxiliem na superação dos obstáculos
9	No que se refere a conseguir fazer o aluno relacionar o conteúdo com outras matérias e assuntos, eu sou nota <u>9,5</u> (1-10)
10	No que se refere a conseguir criar muitas oportunidades para que o aluno perceba a importância daquele conhecimento, eu sou nota <u>9,0</u> (1-10)
11	No que se refere a conseguir levar em consideração, sempre que explico algo, o que o aluno já sabe (ou acha que sabe) sobre o assunto, eu sou nota <u>9,5</u> (1-10)

Questionário 2

6

1	Eu tenho ___% de certeza de que consigo levar em consideração, sempre que explico algo, aquilo que o aluno já sabe (ou acha que sabe) sobre o assunto
2	Quanto a conseguir ajudar o aluno a adotar estratégias de aprendizagem que serão úteis nos próximos anos, eu sou nota ___ (1-10)
3	No que se refere a conseguir ajudar o aluno a identificar as melhores estratégias de aprendizagem, eu sou nota ___ (1-10)
4	Quanto a conseguir fazer o aluno entender como ele esta se saindo na minha matéria, eu sou nota ___ (1-10)
5	Quanto a conseguir fazer o aluno compreender que o uso do material didático é importante para que ele atinja os objetivos, eu sou nota ___ (1-10)
6	No que se refere a conseguir ajudar o aluno que tem mais dificuldade a identificar estratégias de aprendizagem que o auxiliem na superação dos obstáculos, eu sou nota ___ (1-10)
7	No que se refere a conseguir criar oportunidades para que o aluno faça escolhas referentes ao próprio aprendizado, eu sou nota ___ (1-10)
8	Quanto a conseguir criar oportunidades para que o aluno estabeleça seus próprios objetivos de aprendizagem, eu sou nota ___ (1-10)
9	Eu tenho ___% de certeza que consigo fazer o aluno relacionar o conteúdo com outras matérias e assuntos
10	Eu tenho ___% de certeza de que eu consigo criar muitas oportunidades para que o aluno perceba que aquele conhecimento é relevante
11	Quanto a conseguir explicar para o aluno o motivo de cada atividade ser relevante, eu sou nota ___ (1-10)

Appendix C

Signed consent forms (sample)

PARTICIPANT CONSENT FORM

CUREC Approval Reference: Objectives and key results in the classroom: a framework for practice development on goal-setting

		Please initial each row
1	I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. [Eu confirmo que li e entendi as informações referentes ao presente estudo. Tive a oportunidade de considerar as informações e fazer perguntas que foram satisfatoriamente respondidas.]	OK
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or academic penalty. [Eu compreendo que a minha participação é voluntária e que eu sou livre para me retirar a qualquer momento, sem ter que dar nenhuma explicação e sem qualquer consequência ou penalidade.]	OK
3	I understand that research data collected during the study may be looked at by designated individuals from the University of Oxford where it is relevant to my taking part in this study. I give permission for these individuals to access my data. [Estou ciente de que os dados coletados durante o estudo poderão ser examinados por pessoas da Universidade de Oxford onde for relevante. Eu dou permissão para que essas pessoas consultem os meus dados.]	OK
4	I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee. [Eu entendo que esse projeto foi revisado e recebeu autorização do Comitê de Ética da Universidade de Oxford.]	OK
5	I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project. [Eu estou ciente de quem irá consultar os dados pessoais fornecidos, como eles serão armazenados e sobre o que irá acontecer com os dados após o término do projeto.]	OK
6	I understand how this research will be written up and published. [Eu estou ciente de como a pesquisa será escrita e publicada.]	OK
7	I understand how to raise a concern or make a complaint. [Eu entendo como fazer para levantar uma preocupação ou fazer uma reclamação.]	OK
8	I consent to being audio recorded [Eu dou consentimento para ser gravado em áudio.]	OK
9	I understand how audio recordings will be used in research outputs [Eu estou ciente de como as gravações de áudio serão utilizadas na pesquisa.]	OK

10	I agree to take part in the above study. [Eu concordo em participar do estudo.]	ok
11	I agree for my personal data to be kept in a secure database for the purpose of contacting me about future aspects of this study. [Eu concordo que meus dados pessoais fiquem armazenados em uma base de dados segura com o propósito de me contactarem no futuro sobre aspectos dessa pesquisa.]	/

Luiza Silva Soares
Name of participant
[Nome do Participante]

22/02/18
Date
[Data]

Luiza Silva Soares
Signature
[Assinatura]

Name of person taking consent
[Nome da pessoa que pede consentimento]

Date
[Data]

Signature
[Assinatura]

PARTICIPANT CONSENT FORM

CUREC Approval Reference: Objectives and key results in the classroom: a framework for practice development on goal-setting

Please
initial
each row

1	<p>I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.</p> <p>[Eu confirmo que li e entendi as informações referentes ao presente estudo. Tive a oportunidade de considerar as informações e fazer perguntas que foram satisfatoriamente respondidas.]</p>	X
2	<p>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or academic penalty.</p> <p>[Eu compreendo que a minha participação é voluntária e que eu sou livre para me retirar a qualquer momento, sem ter que dar nenhuma explicação e sem qualquer consequência ou penalidade.]</p>	X
3	<p>I understand that research data collected during the study may be looked at by designated individuals from the University of Oxford where it is relevant to my taking part in this study. I give permission for these individuals to access my data.</p> <p>[Estou ciente de que os dados coletados durante o estudo poderão ser examinados por pessoas da Universidade de Oxford onde for relevante. Eu dou permissão para que essas pessoas consultem os meus dados.]</p>	X
4	<p>I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.</p> <p>[Eu entendo que esse projeto foi revisado e recebeu autorização do Comitê de Ética da Universidade de Oxford.]</p>	X
5	<p>I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.</p> <p>[Eu estou ciente de quem irá consultar os dados pessoais fornecidos, como eles serão armazenados e sobre o que irá acontecer com os dados após o término do projeto.]</p>	X
6	<p>I understand how this research will be written up and published.</p> <p>[Eu estou ciente de como a pesquisa será escrita e publicada.]</p>	X
7	<p>I understand how to raise a concern or make a complaint.</p> <p>[Eu entendo como fazer para levantar uma preocupação ou fazer uma reclamação.]</p>	X
8	<p>I consent to being audio recorded</p> <p>[Eu dou consentimento para ser gravado em áudio.]</p>	X
9	<p>I understand how audio recordings will be used in research outputs</p> <p>[Eu estou ciente de como as gravações de áudio serão utilizadas na pesquisa.]</p>	X

10	I agree to take part in the above study. [Eu concordo em participar do estudo.]	X
11	I agree for my personal data to be kept in a secure database for the purpose of contacting me about future aspects of this study. [Eu concordo que meus dados pessoais fiquem armazenados em uma base de dados segura com o propósito de me contactarem no futuro sobre aspectos dessa pesquisa.]	X

Name of participant
[Nome do Participante]

Edson Rocha

Date
[Data]

02.02.18

Signature
[Assinatura]

[Handwritten Signature]

Name of person taking consent
[Nome da pessoa que pede consentimento]

Date
[Data]

Signature
[Assinatura]