THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

EXECUTIVE MSc IN BEHAVIOURAL SCIENCE DISSERTATION

Due Monday 14th December 2020 at 12:00pm

CANDIDATE NUMBER: *(**Five-digit number** available via LSE for You – this is not the same as the number on your LSE ID card!)*	1 8 0 8 0
TITLE:	**Behavioural Change in Organizations: Can Salience Improve the Effectiveness of Institutional Recognition Programmes?**
WORD COUNT: *Excludes references and appendices.*	**9988**
DATE:	

Please tick this box if you would be happy for your dissertation to be shared with future cohorts: ☒

PLAGIARISM DECLARATION:

By submitting this dissertation I certify that I have read and understood the LSE Regulations on Assessment Offences: Plagiarism and formally declare that all work contained within this document is my own apart from properly referenced quotations.

I understand that plagiarism is the use or presentation of any work by others, whether published or not, and can include the work of other candidates. I understand that any quotation from the published or unpublished works of other persons, including other candidates, must be clearly identified as such by being placed inside quotation marks and a full reference to their source must be provided in proper form.

I understand that the electronic version of my work will be submitted to and retained by Turnitin for the purposes of plagiarism detection.

Behavioural Change in Organizations: Can Salience Improve the Effectiveness of Institutional Recognition Programmes?

Abstract

This research explores the effects of recency in improving the effectiveness of an existing recognition programme at a private institution. To create a recency effect, randomly selected staff members within 211 software development teams were exposed to seven internal communications with a reference to the recognition programme, whilst a control group was exposed to the same communications without such device. The experiment led to an 8.3% (p=0.048) marginally significant increase in the adoption of the behaviour promoted by the recognition programme. The effect is reduced to (8.04% p=0.05) after controlling for the covariates identified. The experiment supports the hypothesis that recency could be effective in improving the impact of recognition programmes, thus contributing to accelerate behavioural change in institutional contexts.

Acknowledgements

And whatever you do, whether in word or deed, do it all in the name of the Lord Jesus, giving thanks to God the Father through him.
Colossians 3:17

❊ ❊ ❊

I cannot recall when or where, but it was some time before the start of the programme. Paul was sharing his thoughts about an experiment that involved counting how many calories subjects ate after being primed. He humorously described some colleague of his, named Matteo, who had to inspect the rubbish bins to collect the necessary experimental data while wearing one of those smart Italian suits that characterise him so well. I could never forget this story, because it involved such interesting people who dedicate their lives to the noble cause of doing science to help people have a happier and healthier life.

Approximately eighteen months have passed, and I cannot recognise myself now. I am frequently thinking of formerly obscure terms like *covariates* and *dichotomous variables*. At work, I frequently find myself encouraging colleagues to remember that context matters, or thinking of ways to prime them to get something done with less effort.

This has definitely been the most interesting and fulfilling professional experience of my life, and I am very grateful to everybody who helped me along the way.

I would first like to thank Coco, for bearing with me through so many long hours, and for being so supportive all along.

I think that a good way to measure the impact of our lives is by pondering how many lives each of us have touched, and I want to thank everybody at the programme not for touching mine, but for changing it. Many thanks to Matteo, for making science and statistics interesting, approachable and fun, and for guiding me through the dissertation process. I would also like to thank Paul, Alex, Jet, Therese, Barbara, Umar, Cahal, Christian and everybody else for sharing their knowledge with such sincerity and warmth. May God bless you all.

❊ ❊ ❊

Table of Contents

1. Overview

Recognition programmes have been used in workplaces for more than 40 years. Today, they are among the behavioural management techniques most frequently used. These programmes involve rewarding employees that show desired behaviours with symbolic tokens such as luncheons, gift certificates or verbal expressions of approval, interest and compliments. Recognition programmes have been proven to increase performance at the workplace by 16% on average. However, not all recognition programmes work; they have a probability of success of about 60%.

Recency, a salience effect associated with the tendency to give more weight to the information presented more recently, could be applied to increase the effectiveness of recognition programmes. To date, recency effects are explored and applied in many different fields such as advertising, political communications and energy consumption, with average treatment effects that, to the knowledge of the researcher, vary between 1% and 8%. The application of recency in the context of recognition programmes does not seem to be covered in the extant literature. Rather, the design and communication of these programmes are based on traditional motivation and communication approaches that exploit reflective elements of the mind.

A between-subjects randomised controlled experiment with complete randomisation was performed to assess the effects of introducing recency effects in a recognition programme. Subjects in the treatment group were exposed to various institutional business-as-usual communications over seven weeks, where a salient reference to the newly implemented institutional recognition device was showcased. The control group received the same communications but the reference to the recognition programme was omitted. No other reference to the recognition programme in formal or informal communication or meetings was made during the intervention, in order to preserve the differential recency effect of the treatment.

The recency effect introduced in the experiment produced an average increase of 8.36% (SE=4.21, p=0.048) in the percentage of software tests registered by treated teams. The effect of the treatment was slightly higher (9.64%, p=0.03) among teams with 11 or fewer individuals, which is the ideal team size according to scrum theory. Furthermore, an (9.72% p=0.03) effect was found when excluding teams with atypical size of work scopes, which is reduced to (8.82% p=0.04) after controlling for covariates.

This research extends the academic literature on the empirical effects of salience manifested through recency, which by some accounts may considered limited. The findings support the notion that recency effects could be achieved through routine institutional communications and that recency strategies can be used by practitioners to improve the effectiveness of recognition programmes. This project also offers an innovative measurement approach for assessing the effectiveness of recognition programmes in natural settings. Even if randomised controlled trials are frequently used in fields like marketing, practitioners in fields associated to behavioural management do not use this approach extensively.

2. Literature Review

2.1. The Salience Effect

Our brain has impressive processing capabilities, but it has a limited capacity (Simon, 1955). We cope with the vast amount of stimuli we receive by means of a mechanism of *selective attention* (Hass, 2015). However, even after filtering a large portion of the information we receive, our remaining attention capacity is still limited. Experiments have shown that when we concurrently hear different messages in the left and right ear, we experience important difficulties in paying attention to and recalling one of the two messages. Similarly, our capacity to rehearse a sentence or a sequence of numbers whilst listening to other information is substantially diminished (Broadbent, 1958; Lachter et al., 2004). Such limitations are aggravated by additional declines in cognitive performance due to factors such as stress or even the strains of poverty (Feinberg, 2015, p. 40).

In the context of such limitations, Behavioural Science has found that our attention is powerfully drawn to stimuli that are novel or that seem relevant to us. As described by Taylor and Thompson (1982) "salience refers to the phenomenon that when one's attention is differentially directed to one portion on the environment rather than to others, the information contained in that portion will receive disproportionate weighing in subsequent judgments." In this way, salience induces the observer to select an object or feature and to perform a detailed examination (Neisser & Becklen, 1975). Salience can take many forms, and it is frequently determined by intricate properties of the stimulus such as colour, sound or smell (Dolan et al., 2016; Higgins, 1996, p. 133).

The effects of salience impact our daily lives in various ways. When buying a car, we are more likely to take advantage of automobile price reductions that are directly advertised to us than to price promotions that are exclusively communicated to dealers (Busse et al., 2006) and we are more likely to react to sales tax waivers than income tax credits because the former are more salient due to their immediate and automatic nature (Gallagher & Muehlegger, 2011). When shown two products with equal total cost, we are more likely to choose a product that has a low price and a higher shipping cost, as opposed to one that has a higher price and a lower shipping cost. This might happen because price is easier to encode and compare than shipping charges (Brown et al., 2010; Ellison & Ellison, 2009). We are also influenced by salience in our daily jobs. Financial investors are less attentive to earnings statements on Fridays, when the imminent weekend becomes a powerful distraction (DellaVigna & Pollet, 2009).

Salience effects occur through different psychological processes. Aversion for ambiguity is one of these processes: our attention is drawn to things and situation that we can understand easily, and we avoid choices and situations that imply uncertainty (Alary et al., 2013). Another process relates to looking for anchors to support our decisions about topics we lack knowledge of. In these cases arbitrary and irrelevant numbers that are recently presented to us can bias our judgement (N. Stewart, 2009). Framing ideas and concepts in different ways also alters our choices. Because we do not give the same weight to gains and losses, we process choices differently when they are framed in terms of one or the other (Kahneman & Tversky, 2000). Unusual and extreme experiences are also more prominent in our memories than constant experiences, and such prominence influences the way we make choices (Kahneman et al., 1993). Salience can also offer a reasonable explanation for context dependent preferences and decoy effects (Bordalo et al., 2013; Heath & Chatterjee, 1995) . Recency is another salience effect, and the following pages dwell deeper on it.

2.2. Salience Through Recency

We tend to give more weight to information presented more recently. Our minds are the result of the sum of recent and remote experiences, but not all experiences contribute equally to our current mental state. All other things being equal, the experiences of this morning weigh more heavily than any comparable experience of any morning of one day, one week or one year ago. However, by next month, the advantage of today's experiences will have dissipated and today's events will be barely noticeable (Cherry, 2020). In like manner, a recent experience may momentarily overweigh significant learnings of the past, but this advantage will probably ebb swiftly and the influence of older and more powerful learnings will recover, provided that this recent experience is not atypical or recurring (N. Miller & Campbell, 1959, p. 1). All of these intuitive ideas associated with our brain's decreasing ability to retain information over time are summarized in the negatively accelerated forgetting curve of Ebbinhaus (1913), one of the oldest dependable findings in the field of learning. Of course, our decreasing ability to retain information is altered when an item is consciously reviewed time and again, and if we practice something more than what is usually required to memorize it, the information becomes stored more strongly, reflecting even shallower learning curves (Shrestha, 2017). The discovery of Ebbinhaus enabled additional findings linked to recency. For example, among two associations that are equally strong at any given moment, the older will decay less rapidly (Jost, 1897), and previously reinforced or punished responses can resurface suddenly if the condition is reintroduced (N. E. Miller & Stevenson, 1936).

Recency effects are also associated with the order in which items are presented to an observer. Murdoch (1962) found that items located at the end of a list that has been recently learnt are recalled best, while the first few items are also recalled better than those found in the middle. Such recency effects are possibly the result of our memory processes (Mack et al., 2017). Short-term memory is capable of retaining three to four pieces of information for a brief period (Cowan, 2014; W. H. Edwards, 2010). Another explanation to recency is that memories can form associations with the context in which they are encoded. Temporal and physical contextual cues may help us recall the most recently learnt information, and when such cues are encountered, the associated memories are retrieved (Cohen & Kahana, 2019; Lohnas & Kahana, 2014).

This phenomenon of overweighing recent information has implications for the decisions we make. In a classic experiment by Miller and Campbell (1959), subjects participating as members of a jury were exposed to two arguments, one in favour of the plaintiff and other in favour of the defendant. The jury was up to 5% more likely to remember and vote for whichever argument was presented second. In another experiment involving recency, researchers chose 750 products and placed additional labels that clearly indicated the sales tax that is normally applied at the till. In this manner, they made the total product price prominent when selecting the products, leading to an 8% decline in sales (Chetty et al., 2009).

Marketers also use recency effects to influence the decisions we make. Retention of communication messages is associated with recency, because the repetition of messages can enhance the likelihood of a person retaining the communicated idea (Hovland et al., 1953, p. 10). In political communications, calls to increase voter turnout are more effective when made in the last week of the campaign. Phone contacts boost intention and turnout, but the effect wears off by after a couple of weeks (Nickerson, 2007, pp. 277–279). Online sellers sometimes track the time elapsed since a customer viewed an advertisement and expose the prospective customer to a new message within a short time frame, while the customer may still be in the mood to buy or learn about a product (Spacey, 2017). Similarly, recency can have

an effect on consumers' capability to recall commercial messages, but the average effect found in some studies amounts to as little as 1% (Pieters & Bijmolt, 1997; Terry, 2005).

Energy bills offer another example of recency effects (Hayes et al., 1999). Because consumers receive little feedback on their energy consumption (Darby, 2006), energy bills have been increasingly modified to include sections that show consumers how energy utilisation is spread over time and how consumption changes relative to the prior period or to the same period of the prior year (Hayes et al., 1999). Making such information salient and presenting it continuously has led to a reduction of 4% to 8% of consumption (Sexton, 2015).

The encouraging results of salience in different fields suggest that it can be applied to many novel domains of our daily life. The following section introduces recognition programmes in institutional contexts. Subsequently, an experimental design is proposed to measure the effect of recency in one such context.

2.3. Recognition Programmes: a Novel Field for Salience

Behavioural Management in the Workplace

For more than forty years, organizations have used psychological strategies to modify behaviours in the workplace (Stajkovic & Luthans, 2003, p. 3). These efforts are based on the premise that our actions are a function of contingent consequences and that behaviours that positively affect performance should be reinforced accordingly (Bandura, 1969; Komaki et al., 1996). Behavioural Management efforts are rooted in early ideas proposed in Skinner's Operant Conditioning (2019) and Thorndike's Law of Effects (1913), which assume that reinforcers are the causal agents of behaviour (J. Komaki et al., 1996). There are three traditional approaches to attempt to modify behaviour: contingently administered money, feedback, and social recognition (Bandura, 1986; Mitchell & Mickel, 1999; Stajkovic & Luthans, 2001).

Recognition programmes offer symbolic rewards to employees that exhibit the desired behaviours. Traditional recognition programmes use luncheons, gift certificates, or personal attention through the use of verbal actions such as expressions of approval or compliments (K. Luthans, 2000, p. 34; Stajkovic & Luthans, 2003, p. 10). The employees' responses to recognition can be associated with motivations such as reciprocity, altruism towards employers who care for their employees, or desires to conform with internal group norms (Bradler et al., 2016). It has been theorised that social recognition at work derives motivation from its predictive value, because desired personal consequences such as work promotions or raises are usually preceded by social approvals (Luthans & Stajkovic, 2000). Meta-studies measuring the effects of behavioural management suggest that recognition programmes, when implemented without additional feedback or contingent monetary rewards, improve performance by 16% on average (Stajkovic & Luthans, 2003, p. 23)

Recency in Recognition Programmes

The effects of recency on the effectiveness of recognition programmes is an current gap in the literature. Until now, the communication of recognition initiatives has been mainly designed using traditional approaches and the extant literature offers classical considerations such as the clarity of the message, its frequency and the selection of the adequate

communication channels (Quirke, 2012) (Weik, 1995). Different queries were performed in the main literature search engines to identify salience techniques applied to internal communications and to recognition programmes, but the results were ineffectual. It seems that the existing literature fails to account for the effects of salience strategies on recognition programmes. Whether salience can increase the effectiveness of behavioural management initiatives or contribute to recognition programmes in particular is a relevant question that we believe remains unanswered.

The following pages outline a novel experiment to assess the effectiveness of recency on a recognition programme. To achieve this, a treatment group receives institutional communications that make a salient reference to such recognition programme, whilst a control group receives the same institutional communications without any reference to such salience device.

3. Experimental Context, Research Questions, Hypotheses and Variables

3.1. Research Objectives

This research explores if salience can be applied to a recently implemented recognition programme in order to improve its effectiveness. This exploration contributes to filling a gap in the literature and enriching the body of knowledge of salience in the context of recognition programmes, particularly within institutional contexts. It may also shed light on further applications of salience to other behavioural management approaches, or to similar institutional settings where adoption of new behaviours is required.

3.2. The Context of the Experiment

An Institution Adopting Agile Software Development Practices

This experiment takes place in a natural context where software development teams code and implement new software systems within a private institution. The institution performs software development projects ($n \sim 300$), each of which is carried out by a full-time project team. The terms “project”, “team”, “project team” and “scrum” will be used interchangeably to refer to a scrum team: a group of 2 to 10 individuals that work exclusively and full-time to produce a previously-agreed software product. Three different roles exist within each team. The *product owner* defines the project *features* and prioritises the product backlog. The *scrum master* is responsible for promoting adherence to scrum methodology and for resolving the obstacles that the team faces as it carries out the project. The remaining members are called *“team members”*. They are responsible for developing, testing and launching the software product.

The institution is investing to institutionalise *Agile* practices among its software development teams. *Agile* is a software development approach based on the Japanese Kaizen philosophy; it operates through a working paradigm that differs significantly from western software development practices (Conboy et al., 2011; Kompella, 2014). For example, western firms follow the practice of performing extensive planning through Gantt charts, whilst *Agile* promotes “lightweight planning” through simple work backlogs and extensive use of kanban boards to track progress (Hesselberg, 2018; Salameh, 2014).

The experiment will focus on the adoption of the practice of registering and tracking software test. All scrum teams are expected to register at least one software testing activity for each committed feature. These tasks are registered in a software system used to track progress. To encourage teams to perform software tests, the institution routinely sends scrum masters different email communications with tips, recommendations and reminders.

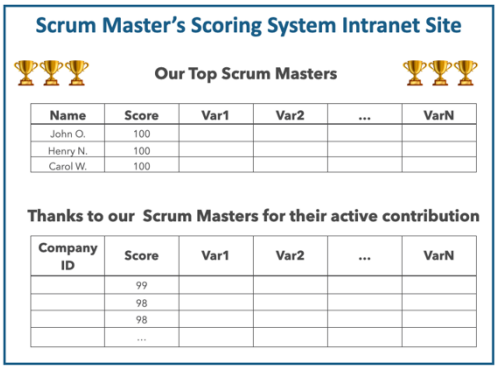
A Recently Implemented Recognition Programme

The institution employs a performance scorecard to measure, among scrum masters, the adoption of behaviours associated with *Agile*. Ten different variables such as: training in *Agile*; training in using the institutional kanban boards; and adequate registration of features are considered. Each variable accounts for a number of marks that add up to a total of 100 possible marks. Scores are kept for every quarterly cycle and are reset to zero when a new quarter starts.

Past and current scores are published in an intranet site for scrum masters to review. Recording and performing software-testing activities for each committed feature is a variable considered in the score and is equivalent to 10 marks.

The Researcher recently implemented a recognition programme connected to the performance scorecard. The recognition programme offers the status of "Top Scrum Masters" to those individuals who gain high marks in the institutional scrum master's performance scoring system. This recognition device seeks to activate improvement motives among scrum masters through our tendency to "self-evaluate" (Tesser, 1986, 1988) and offers a positive social identity to those who gain the status of "Top Scrum Masters" (Tajfel et al., 1979). Scrum masters with 100 marks scores are named at the top of the intranet site and praised as "Top Scrum Masters," as shown in **Figure 1**. They also receive a formal recognition once every quarter, which is delivered by a Senior Manager from the institution. Scrum masters which obtain a lower score can verify their own score by searching their employee number rather than their name, in order to avoid shaming those with low scores. During the third quarter of 2020, 17 scrum masters out of approximately 200 obtained the "Top Scrum Master" accolade.

Figure 1: Current Display of the Scrum Masters' Performance Score



3.3. Research Question and Hypotheses

The research would like to answer what is the difference in the adoption of a novel work practice when a recognition device is made salient in routine communications. A detailed research question and null hypothesis is presented below.

Table 1: Research Question and Hypotheses

Research Question	What is the difference in the percentage of features that have a registered software testing tasks between those teams where the scrum master has received routine institutional communications with a salient recognition device and those teams where the scrum master has not been exposed to such device?
Null Hypotheses	There **is no** difference in the percentage of features that have registered software testing tasks between those teams where the scrum master has received routine institutional communications with a salient recognition device and those teams where the scrum master has not been exposed to such device.

3.4. Variables

The researcher proposes the following experimental variables:

Table 2: Proposed Experimental Variables

Dependent variable	The extent to which a team registers and performs software testing tasks is measured by: • **Y:** Percentage of features within each scrum backlog that have one or more registered software tests in the institutional tracking system.
Independent variable	The independent variable proposed reflects the delivery of the treatment: • **X:** Exposure to routine communications where the institutional recognition device is made salient.
Scrum characteristics for correlation analysis	Other variables will be considered in the experimental design but may still be correlated with the dependent variables, as they vary across teams: • **SMT:** Whether the scrum master has received Agile training. • **SMN:** Whether the scrum master is new to the team. • **TS:** Number of individuals in the scrum team. • **CF:** Number of features committed by the scrum.

4. Methodology

4.1. Experimental Design

Different aspects of the internal organisation of the institution are relevant to the experimental design. The institution traditionally carries out approximately 300 active projects sponsored and funded by one of thirteen different departments. A portfolio manager leads each department's projects basket and sets the general strategy and priorities. Departments are relatively isolated from each other; some work in separate buildings while others do so on different floors within the same premises. They have influential figureheads and relatively well-knit sub-cultures.

Each department's projects basket is organized into smaller groupings called programmes, led by a programme manager who is responsible for tactical and operational matters. Programmes are smaller units where there is a robust internal cohesion promoted by the programme manager. Within programmes, and to a lesser degree within departments, there are robust informal networks characterised by fluid communication. Consequently, treatments are more likely to spread to other scrums within a programme than to other scrums within a department.

Considering the above, three main options for the levels of randomisation were available in the experimental context:

- ***Department level:*** Each department could be allocated to a treatment group. This experimental design offers the advantage of a reduced risk of spillovers. Conversely, this design carries a high risk of confounding factors affecting results because of departmental fixed effects.
- ***Programme level:*** Within each department, programmes could be randomly allocated to different treatments and all the scrums within each programme would receive the same treatment. This randomisation level offers a lower risk of confounding factors and spillovers.
- ***Project level:*** Each project team can be randomly allocated to either treatment or control groups and all members could receive treatment if allocated to the treatment group. This configuration carries a higher risk of spillovers because of the porosity of informal networks within programmes and departments.
- ***An intervention addressed exclusively to scrum masters:*** Each scrum team has a full-time individual appointed to the scrum master role. Addressing the intervention solely to the scrum masters could reduce the risk of spillovers, as fewer individuals within the ecosystem would be aware of the treatment. Additional information is presented in **Appendix 1.**

Cluster Randomised Controlled Field Experiment Design

Due to the porosity of the institutional social networks, randomising at the scrum level was perceived to increase the risk of spillovers. A cluster randomised controlled field experiment at the programme level would have offered the possibility of performing this intervention when it is not desirable to randomise at the scrum level (S. J. L. Edwards et al., 1999). Randomising at the programme level reduces possible resentment effects or feelings of unfairness as fewer untreated individuals are likely to be affected by spillovers than in other configurations. Such design would reduce the risk of confusing factors influencing the treatment effect because of the individual characteristics of each department. If properly designed and executed, this design would meet assumptions of excludability and non-interference and would be internally valid, offering an unbiased estimate the average treatment effect (Galizzi, 2020).

However, this experimental design offered relevant risks to internal validity. Cluster randomisations produce a more substantial sampling variability when compared to complete randomisation and tends to require a larger sample to obtain the same statistical power (Campbell et al., 2004). Another problem of this experimental design is intra-cluster correlation caused by the similar profiles of individuals within clusters, which requires additional calculations to estimate interclass correlations, because failures to account for it will negatively affect statistical power and may induce to Type-1 errors (Murray et al., 2020).

Considering the above, the researcher performed a preliminary sample size calculation to assess the viability of such configuration design, which revealed that the universe of circa 80 programmes would offer a statistical power that could not detect an ATE smaller than 7%, which the researcher deemed too daring at the planning stage. Details are presented in **Appendix 2**. **Table 3** considers additional scenarios performed through a sensitivity analysis, but none seemed feasible during the planning stage.

Table 3. Sensitivity Analysis of Sample Size Using Different Parameters and α=0.05

Scenario	Mu1	Mu2	Sd1	Sd2	M	Rho	Ind. Rand.	Sample per arm	Clusters per arm
To detect an ATE=5; Low ICC.	50	54	11	11	4	0.3	119	232	58
To detect an ATE=5; Med ICC.	50	54	11	11	4	0.5	119	304	76
To detect an ATE=5; High ICC.	50	54	11	11	4	0.9	119	444	111
To detect an ATE=7; High ICC.	50	57	11	11	4	0.9	39	147	37
To detect an ATE=7; Med ICC.	50	57	11	11	4	0.5	39	104	26

Between-Subjects Randomised Controlled Experiment with Complete Randomisation

A between-subjects randomised controlled experiment with complete randomisation targeted solely at scrum masters was chosen as a second-best alternative to safeguard non-interference (Haynes et al., 2012). Under this configuration, the researcher randomly selected treatment and control samples of project scrum masters that fulfil a set of minimum criteria, as explained below. To ensure both groups are statistically similar, they were compared with respect to the measures expressed in the previous section (Glennerster & Takavarasha, 2013, p. 98)

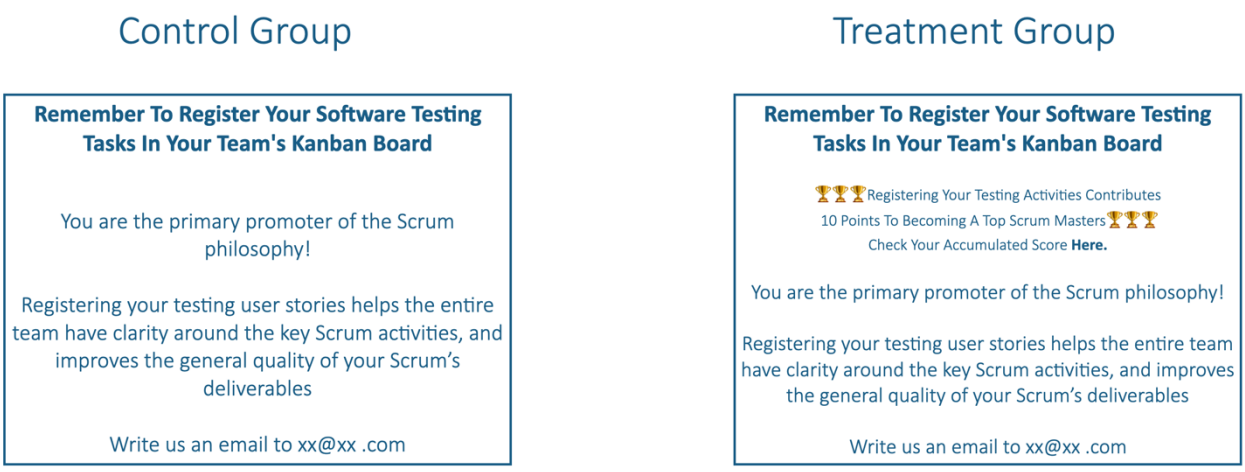
The chosen experimental configuration offers at least four advantages. Randomised experimental designs yield the most accurate estimations of the effect of an intervention. If properly designed and executed, such experiments are internally valid and provide an unbiased estimate of the average treatment effect (Galizzi, 2020). Furthermore, spillovers are less likely to materialise than in an experimental arrangement where the entire scrum team receives treatment, because a smaller proportion of individuals within the ecosystem would receive treatment. This design would also reduce the risk of confounding factors occurring due to possible correlations between the treatment effect and the individual or temporal characteristics of any department or programme. Further, because of the natural setting, this experiment offers an enhanced external validity, as the project teams and institutional rules are part of a real-world environment (Harrison & List, 2004).

However, the selected experimental design offers certain drawbacks that must be understood. Firstly, because the treatment is addressed exclusively to the scrum master, it could be less effective than a treatment directed at the entire scrum team. Secondly, violating the SUTVA assumption through informational spillovers to untreated scrum masters, though less likely, are still possible. Spillovers could materialise either through the informal commentary of the intervention or by gaining access to the communications where the treatment is exposed. The impact of such a violation will consist of a positive spillover that could bias the treatment effect downward. This would increase the probability of an effect happening but not being detected because Control Group individuals could receive the treatment. Thirdly, the proposed experimental design could also trigger resentment effects or feelings of unfairness among untreated individuals. To counter this risk the researcher was prepared to explain that the treatment was being delivered in this quarter to a group of scrum masters, but that the remaining scrum masters would receive the treatment in the following quarter, so that everyone could participate in the initiative (Glennerster & Takavarasha, 2013, pp. 317–321).

4.2. Manipulations Performed

As explained above, the institution performs routine communications to scrum masters proving guidelines on how to register and perform software tests. A Treatment group received seven communications with a salient reference to the newly implemented recognition device, in an attempt to activate a recency effect, whilst a control group received the same communications without the recency device. A scheme of the salient reference to the recognition programme is shown in **Figure 2**. The intervention ran for seven weeks between 1 October 2020 and 19 November 2020.

Figure 2: Simplified version of one communication addressed at Treatment group as part of the experiment



4.3. Ethical Considerations

The experiment can raise ethical considerations because the scrum masters participating in the experiment will not be aware of the intervention. A full ethics review form was submitted to the University's Research Ethics Committee and is enclosed in **Appendix 3**. The Research Ethics Committee's approval is presented **Appendix 4**. The Committee requested that participants involved in the study receive a de-brief or post-study information. This will be performed in January 2021.

4.4. Participants

Treatment and control groups of scrum masters were selected among the scrum teams that fulfil the following characteristics:

Table 4: Sample Characteristics

Sample Characteristics	• Active scrum during 4th quarter of 2020 • Appointed scrum master • Scrum and Kanban Teams • Scrum master email available • At least one Feature registered in the quarterly backlog

Selecting scrums with these characteristics offers strengths but at the same time produces certain limitations. One strength is that these selection guidelines do not impose many restrictions on the team's eligibility criteria, helping increase the sample size. Another strength of this configuration can increase external validity to the study because scrums with different traits are more representative of scrum teams in any professional practice. Conversely, one disadvantage of this configuration is that accepting teams with very large team sizes, or with many features committed may introduce additional factors that may increase the variability of results, reducing power This is explained in detail in the Results and Discussion sections.

Controls to avoid selection bias were necessary for this step. To prevent selecting participants that share a common characteristic that could affect the results, a member of the Metrics Collection Team at the institution identified the scrum teams that fulfil the abovementioned characteristics without having any knowledge of the experiment. Afterwards, she shared with the researcher a coded list of scrums that meet the abovementioned characteristics. The researcher then carried out the randomisation of teams, as explained below. Because of its coding, the researcher was not be able to discern any information about the scrums in the list.

4.5. Sample Size

The expected treatment effect was obtained from prior experiments, and the standard deviation was obtained from the natural context where the experiment takes place. Recognition programmes, when used individually (as is the case of this study) increase performance in 16% on average with a probability of success of 60% (F. Luthans & Stajkovic, 2000, p. 27). Such parameters offer an expected value of 10.71%. In a separate salience study, Chetty et al. (2009) found that salience produced an average treatment effect of 8%, which will be assumed for this study as an estimate. Actual compliance with the registration of software tests was assumed at 50%. Combining the assumed current compliance of software test registration and a treatment effect of 8%, the expected impact of the treatment was assumed at 4% (0.5*0.08=0.04). Regarding the standard deviation, during the third quarter of 2020, teams correctly registered 93% of work petitions to internal support areas with a 10.7% standard deviation. This was the only available estimation of a possible standard deviation associated to the registry of scrum software tests. Combining these parameters, a sample size of 119 teams for treatment and 119 for control would successfully detect a 4% difference in the correctly registered support petitions (using Stata: sampsi 50 54, sd[11] alpha[0.05] power[0.8]). These samples fell within to the population size of c. 250 projects expected to be active in Q4-20 during the planning stage.

4.6. Randomisation Procedure

To decrease the risk of sampling bias, randomiser was used to perform the necessary randomisations. The list of scrum teams that fulfilled the minimum characteristics explained above was firstly randomised through a random sequence. After they were randomly organized, a random assignment was performed using that same tool.

4.7. Data Collection and Analysis

A reliable internal data collection process exists at the institution where the study was performed. Internal processes track the scrums' performance at the team, programme and department level, and they produce different performance data. Baseline data on all the variables explained above will be collected through these processes.

To safeguard excludability, the members of the Metrics Team that collected data were not aware of which scrums were assigned to the Treatment or Control Groups. In like manner, the researcher was not involved in the collection of results.

After gaining access to the data collected by the metrics team, the researcher performed the following analyses:

- A review of the descriptive statistics of the dependent variable in general sample, and in the control and treatment group.
- A graphical inspection of the data.
- A pairwise correlation of the dependent variable with the exogenous variables mentioned in **Table 2**.
- A normality test of the dependent variable using Shapiro-Wilk.
- A parametric hypothesis test, as evidence of non-normality was not obtained.
- Hypotheses tests to inspect the robustness of the results among subsamples that reflect teams' maturity, size and work scope.
- General and multiple linear regressions considering relevant covariates identified in the pairwise analysis.
- An *a posteriori* sample size estimation.

5. Results

5.1. Treatment and Control Group Composition

The Stata do file used to analyse the data is presented in **Appendix 5**, and **Appendix 6** shows the detailed statistical outputs produced. The dataset is comprised of 232 active scrum masters that fulfilled the minimum characteristics. 124 (53%) scrum masters were randomly allocated to the Control Group, and 108 (47%) to the treatment Group.

5.2. Attrition

Attrition amounted to a total of 21 scrum masters, which represents an 8.6 % of total teams originally considered in the sample. One project had a change of scrum master during the quarter, four projects were unexpectedly cancelled, and sixteen teams declared that no testing activities were to be performed because of particular characteristics in their scope or approach These teams were identified through routine business-as-usual interactions between scrum masters and the measurement team. After considering attrition, treatment and control groups were tested to verify that they are well balanced in seven key characteristics, as shown below.

Table 5: Scrum Characteristics Across the Control and Treatment Group

	Shapiro-Wilk Prob > z	*Test applied*	*Prob > |z|*
Project importance rank	0.00	Mann-Whitney	0.88
Mix of new vs ongoing projects	0.00	Chi2	0.38
Scrum master performance score in prior quarter	0.00	Mann-Whitney	0.51
Number of features committed	0.00	Mann-Whitney	0.7
Team size	0.00	Mann-Whitney	0.75
Continuing vs new scrum master in the team	0.00	Chi2	0.55
Scrum master training	0.00	Chi2	0.45

5.3. Descriptive Statistics

The percentage of features with a software test task among the entire sample amounted to Y = 57.9% with median= 0.60 and SD=30.6%. Y_T exhibits fewer extreme events than a standard normal distribution (K = 1.99) and is approximately symmetric (Sk = -0.16).

The treatment group average is 8.37% higher than the Control Group (Y_T = 62.50% versus Y_C = 54.13%) and exhibits a higher median (Mdn_T = 66.67%; Mdn_C = 50%). Treated subjects exhibits a lower standard deviation (SD_T = 29.66% versus STD_C = 31.03%), are more skewed to the left, reflecting a larger percentage of features with a recorded test (Sk_T = -0.27; Sk_C = -0.07), and have a similar kurtosis (K_T = 2.02; K_C = 1.99).

5.4. General Test of the Null Hypothesis

The salience effect introduced in the experiment produced a marginally statistically significant effect within the treatment and control groups considering an α of 0.05. To test the hypotheses, the researcher could not reject that Y comes from a normal distribution using a Shapiro-Wilk test (W=0.99, p=0.71), and a standard parametric test was performed for the dependent variable. **Table 6** suggests that assuming normality and standard assumptions, the difference of *Control* and *Treatment* means [diff=-8.36; SE=4.20] is statistically significant using a two-tailed test. Based on this test, the researcher rejects the null hypothesis.

Table 6: Parametric Test of the Null Hypothesis

Group	*Obs.*	*Mean*	*Std.Error*	*Std. Deviation*	*95% Conf. Interval*
Control	116	54.14	2.88	31.03	48.27 59.84
Treatment	95	62.50	3.04	29.65	56.45 68.54
Combined	221	57.90	2.11	30.63	53.74 62.06
Difference		-8.36	4.21	--	-16.66 -0.063
Results					T = -1.98 Degrees of freedom= 209 P>|t|=0.0483

Basic and Multiple Linear Regressions

For the complete sample, a basic linear regression is a better fit that an intercept-only regression model [F(1,209)=3.95; Prob>F=0.048], but only accounts for 1% of the behaviour of the dependent variable (Adj R_2=0.01). This model estimates that the treatment increased the percentage of features with test in 8.36% (SE= 4.21 P>|t|=0.048) with a vertical intercept $\boldsymbol{\beta}_1$=54.13% (SE=2.82, P>|t|=0).

The institution collects data on certain exogenous variables that influence the performance of projects. These variables include the projects' importance (measured by its rank in the general projects list performed every quarter), whether the project is new or ongoing, whether the scrum master has received formal training, the team size and number of committed features, and whether the scrum masters is new to the team or not. A pairwise correlation (shown in **Appendix 6** revealed that the total number of features registered by the scrums is the only exogenous variable that causes a statistically significant covariate effect ($\boldsymbol{\rho}$ =-26%, α=0.05) with the dependent variable. A multiple linear regression considering the number of total features registered produces a better fit that an intercept-only model [F(2,208)=9.61; Prob>F=0.00], and the combined variables account for 7.58% of the behaviour of Y (Adj R_2=0.08). The vertical intercept, *number of features* slope and *treatment* slope are statistically significant. This model predicts a 72% intercept (SE= 5.37; P>|t|=0.00) and -3.53% (SE= 0.91; P>|t|=0.00) fewer software tests for every feature registered by a Team. The treatment effect is estimated in $\boldsymbol{\beta}_T$=8.04 (SE=4-08; P>|t|=0.050).

5.5. Robustness of Results Among Different Subsamples

The researcher does not propose a multiple hypothesis testing correction for subsample analyses (Dunn, 1961). However, a programmed replication discussed below has preregistered different hypotheses for the subsamples. **Table 7** shows that the treatment only had a statistically significant effect on the teams that have 11 or fewer members.

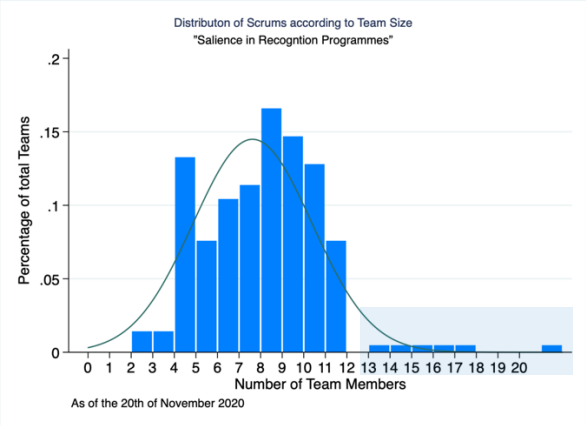
Table 7: Robustness of Results Among Relevant Subsamples (Including Outliers)

Dimension	*Subsample*	*Difference*	*SE*	*P value*
Team Maturity	Ongoing Teams	-6.13	4.58	0.18
	New Teams	-20.43	10.82	0.07
	Ongoing Scrum Master	-6.92	4.51	0.13
	New Scrum Master	-17.12	11.96	0.16
Team Size	Teams with 11 or fewer members	-9.65	4.28	***0.03***
	Teams with more than 12 or more members	N/A	N/A	N/A
Committed Scope	Teams with 7 or fewer Features	-7.17	4.46	0.11
	Teams with 8 or more Features	-14.74	11.67	0.22

5.6. Robustness of Results Excluding Outliers

Agile theory emphasizes that team size and committed work scope have an important effect on performance (Progress Corp., 2013; Scaled Agile Inc, 2019). **Graph 1** presents the distribution of teams according to team size and number of committed features and suggests that six teams with more than 11 individuals and one team with 15 registered features could be considered as outliers within the sample.

Graph 1 : Distribution of Teams According to Team Size and Number of Features



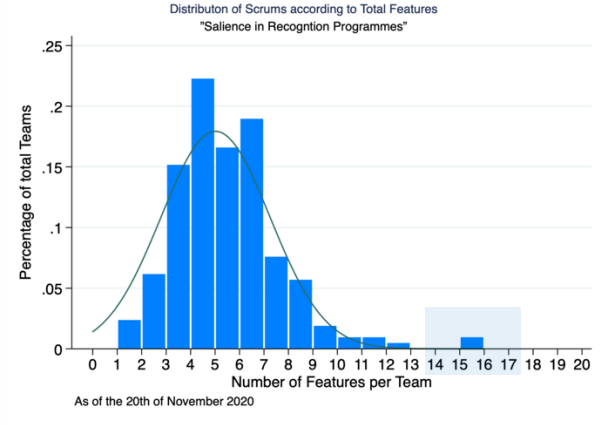


Table 8 shows that assuming normality and standard assumptions, the difference of the control and treatment means excluding outliers (diff=-9.72; SE=2.16; p=0.02) is statistically significant using a two-tailed test, offering evidence to reject the null hypothesis among the subsample that excluded outliers.

Table 8: Parametric Test of the Null Hypothesis Excluding Outliers

Group	*Obs.*	*Mean*	*Std.Error*	*Std. Deviation*	*95% Conf. Interval*
Control	114	53.66	2.90	31.00	47.91 59.41
Treatment	90	63.38	3.15	29.84	57.13 69.63
Combined	204	57.95	2.15	30.80	57.70 62.20
Difference		-9.72	4.30	--	-18.20 -1.25
Results					T = -2.26 Degrees of freedom= 202 P>|t|=0.02

When excluding outliers, a basic linear regression is also a better fit that an intercept-only regression model [F(1,202)=5.12; Prob>F=0.02], but only accounts for 2% (Adj R_2=0.02) of the behaviour of the dependent variable. This model estimates that the treatment increased the percentage of features with test by 9.72% (SE= 4.3; P>|t|=0.03) with a vertical intercept $\boldsymbol{\beta}_1$=53.7 (SE= 2.9; P>|t|=0.00). Controlling for the effect of the total number of features registered by scrums [F(1,201)=11.06; Prob>F=0.00], the combines variables account for 9% of the behaviour of Y_T (Adj R_2=0.09). The vertical intercept, *total features* and *treatment* slopes are all statistically significant. This model predicts an average treatment effect of 8.82% (SE= 4.1 P>|t|=0.04)

6. Discussion

6.1. Average Treatment Effect

The treatment produced a marginally significant effect among the general sample, and stronger effects were found in some subsamples. The results reflect a statistically significant increase in the percentage of registered software tests of 8.36% (SE=4.21; p=0.048) among teams treated. This result is extremely close to the cut off level of 0.05 but is still significant. The treatment effect lies within the upper range of prior salience effects reported in the literature, which range between 1% and 8%, and is below the 16% average improvement in performance reported when using recognition programmes in institutional settings. The treatment produced a slightly higher average effect of 9.65 (SE=4.28 p=0.03) among teams that have 11 or fewer individuals. This is consistent with scrum theory, which suggests that teams with such dimension are more effective. In like manner, when excluding outliers in terms of team size and committed work scope, the treatment effect increases to 9.72% (SE=4.3; p=0.02). This is also consistent with theoretical propositions that teams are more effective when they commit moderate work backlogs.

The results are also encouraging after controlling for the effect of covariates. When considering the effect of the total number of features registered by the scrums, the average treatment effect adjusts to $\boldsymbol{\beta}_T$=8.04 (SE=-4.08, P>|t|=0.050) for the entire sample. When outliers are excluded, a higher average treatment effect of 8.81% (SE= 4.1 P>|t|=0.04) is observed.

Despite these results, the robustness of the treatment results does not extend to three relevant subgroups within the general population, and more research would help to gauge if recency can be effective in such spaces. Ongoing teams frequently poses greater clarity about the work within scope and could be expected to reflect a stronger response to invitations to record software tests (Diff=-6.13, p=0.18). In like manner, scrum masters that are not new to a team may have above-average knowledge of the work to be performed, and thus would be expected to exhibit a clear response to the treatment (Diff=-6.92, p=0.13). Furthermore, the organization encourages teams to commit no more than seven features on any given quarter, because as explained above, a moderate work scope is associated with an improved execution of work. However, no results were observed in this subsample (Diff=-7.17, p=0.11).

The recent introduction of the recognition programme might have subdued the general treatment effect and might explain the lack of results among important subgroups. In its managerial role, the researcher introduced the device in the institution in September of 2020. Though scrum masters have been formally made aware of its existence, the device has not received an extensive dissemination throughout the institution. For example, at least 10 scrum masters approached the institutional team responsible for calculating the scorecard with different doubts about its calculation.

6.2. Validity

The research sought to achieve internal validity through several means. A randomised controlled trial was proposed because it is the one of the most rigorous and robust research approaches for determining an unbiased estimate of the treatment. A randomisation at the scrum master level attempted to reduce positive spillovers by drastically reducing the density of individuals treated in the ecosystem. To prevent favouritism in the sample selection, the list of possible participants was prepared by the institution without providing the researcher specific details of the teams involved. The

research used a reputable randomisation software available online to avoid sampling bias and the treatment was delivered by an individual not involved in the randomisation (Ards et al., 1998). To seek double blindness, care was taken to ensure the individuals responsible for measurement processes at the firm did not have knowledge of who the treated individuals were. Finally, in a quest to avoid Hawthorne effects, the individuals participating in the experiment had no knowledge of it, and authorization from the LSE and the institution was sought accordingly (Gimotty et al., 2002).

At the same time, the natural context of the experiment may offer external validity to the findings reported. The sample used in this experiment encompasses teams (n=211) and individuals (n=1716) with very different characteristics. Such context is rich in real-world situation where subjects, stimuli, tasks, rules and stakes combine to determine the results obtained (Harrison & List, 2004, p. 1012). For this reason, the setting of the experiment suggests that its findings could have implications across situations with similar stimuli, individuals, and contexts, particularly when such results are pondered in connection with related scientific findings (Calder et al., 1982). Future replications across different setting will hopefully offer addition information in this regard (Goodman et al., 2016).

6.3. Research Contributions

This research extends the academic literature on the empirical effects of salience manifested through recency. This may be relevant because some researchers argue that the body of evidence supporting the applicability of behavioural interventions across a variety of real world-contexts is still relatively thin (Bhanot et al., 2018, p. 1). To this date, recency effects been identified and measured in experimental and natural contexts covering different fields. This research provides additional field evidence on the effectiveness of recency, and may be used to complement the existing literature (Harrison & List, 2004).

Secondly, this research supports the hypothesis that recency can be useful to improve the effectiveness of recognition programmes, a theme that does not appear to be covered in the literature. To this date, mainstream recognition initiatives appear to be designed with the perspective of traditional motivation and communication theories, which deeply rely on reflective processes (Latham, 2012). Conversely, this research proposes that simple environmental alterations can activate recency processes which are automatic in nature, and that these automatic processes may improve the effectiveness of recognition programmes (Dolan et al., 2012; Vlaev & Dolan, 2009).

Furthermore, this research suggests that recency effects could be achieved through routine institutional communications that follow similar guidelines. To maintain the task of registering and performing software tests in the memory of scrum masters, the experiment relied on disseminating a few routine institutional communications that covered different topics related to software tests. Certain messages specifically invited scrum masters to register and perform software tests, whilst other messages offered guidance on additional tasks associated with software quality. Such complementary messages are more likely to lead to a positive evaluation of the message (McCullough & Ostrom, 1974) Secondly, no more than one message was sent every week. Even if the repeated presentation of persuasive messages can increase retention, the impact of the message can be reduced due to excessive repetition (Appel, 1971; Cacioppo & Petty, 1980).

6.4. Implications for Policy Makers

An Effective Behavioural Tool

This research offers evidence that practitioners can use the low-cost behavioural effect of recency, in order to improve the effectiveness of recognition programmes. Institutions invest heavily to changes the way people behave by attempting to change their minds, but these investments frequently fail (Dobbin & Kalev, 2016; Morse, 2016). This happens because changing minds requires that our cognition transit several stages (Kolb, 1984), with a particular mindset (Dweck, 2006). In addition to their low effectiveness, efforts to change behaviours that rely on reflective processes have additional disadvantages. They can increase inequalities by benefitting those who are better prepared to assimilate these treatments (Lorenc et al., 2013), they can be very expensive because they are labour intensive and require extensive contextualisation (Schippers et al., 2015) and they frequently produced unexpected and undesired effects(Petrosino et al., 2000). In contrast, recency is an inexpensive, simple and scalable behavioural strategy that can bring high impact changes (Banerjee et al., 2011; List & Gneezy, 2014; Thaler & Sunstein, 2009).

A Reliable Mechanism to Measure Behavioural Change in a Natural Context

Despite the long history of recognition strategies, measurement their impact in natural contexts is still difficult. In their attempts to modify behaviours, firms typically deploy mutually reinforcing efforts such as economic incentives, feedback processes recognition programmes, or even training and alternative work structures (Stajkovic & Luthans, 1997). Efforts are frequently directed from different departments; they often occur simultaneously and in some occasions without proper coordination between the involved areas (Moran & Brightman, 2000, p. 69). In such conditions, obtaining an unbiases estimate of the effect of recognition programmes is still a challenge for practitioners (Daunt & Menzies, 2020).

This research offers an innovative means to measuring behavioural change in institutional contexts. Randomised controlled experiments offer a practical method to isolates the effect of recognition devices used in organisations (Galizzi, 2020). The approach followed in this document helps to discern unbiased estimates of a recognition intervention, by taking into account the effects of external factors, selection bias, and other threats that hinder an accurate measurement. Today, experimental designs to identify average treatment effects are relatively common in certain organisational domains like marketing (Gallo, 2017). However, such designs are not as common among human resources or behavioural management practitioners (Haak, 2013).

6.5. Limitations

The design of the experiment was insufficiently powered to detect the effect of the Treatment across the entire sample. The original experimental design considered 119 groups for each treatment and control groups. Due to the economic contraction caused by the pandemic, the total number of projects that fulfilled the minimum conditions during the fourth quarter of 2020 was reduced to 232. In addition, the experiment was designed to detect an 8% effect assuming a Standard Deviation of 11%. In contrast, the standard deviation in the sample was 31%. An *a posteriori* sample size reveals that the minimum sample size to detect an effect of 8.36%, (as presented in the general test of the null hypothesis) is n= 207 for each control and treatment groups.

Many important elements that influence the performance of scrums were not accounted for, and other variables considered in the study only partially reflect the elements they attempt to measure. On average, scrums are active during four quarters, but there is a great dispersion in their average lifespan. The teams' stage in their lifetime has implications in terms of stability and performance. The study only considered a dichotomous variable that records whether teams are new or ongoing, but more accurate data on the number of prior active quarters could have improved the subsamples analysis, and it could have dilucidated other possible covariates and their effects on the treatment. Training among teams another example: the study only accounts for whether the scrum master received training, but no information is offered about the team's training.

Furthermore, the intervention failed to capture the final three weeks of the project cycle, a period where a large part of software testing activities are recorded in the institutional system. Quarterly projects cycles are subdivided into six two-weeks periods. The first two-weeks period of the fourth quarter (25 September to 2 December) were dedicated to the project set up. The following four weeks (5 October to 27 November) were dedicated to project execution. The final two-week period 30 November to 11 December) were dedicated to carry out project closures activities and to planning the next quarter. This experiment concluded on the 20th of November and was shortened by three weeks in order to align its schedule with the research submission dates. Ideally, the intervention should have extended until the 11th of December, which is the date when the project execution cycle ends. Such extension would have been naturally aligned to natural context of the experiment and could have helped to capture and measure the software test activities that are registered towards the end of the quarter.

Finally, this research did not account for the effect of different degrees of engagement with the treatment. Communication efforts frequently measure awareness of messages by identifying the number of emails opened. Engagement can be complemented by measuring for example, the number of different individuals that actually reviewed their current results in the webpage where the internal recognition programme is presented (Wells, 2017). No data was collected on engagement because blind carbon copying, which was used for distribution of the treatment and control emails, does not allow read-receipt functionality, and a manual follow up would have been costly and time consuming. Employees may have acquired different degrees of awareness of the treatment, depending on the number of emails they actually engaged with, and modelling the differential treatment effects in terms of the intensity of the interaction with the emails would have been valuable.

6.6. Future Research

Based on the encouraging findings of this study, a programmed replication of this experiment will run from December to March 2021 with a number of improvements. A pre-registration is currently awaiting approval, and a copy of the document has been uploaded with the dissertation documents (S. Stewart et al., 2020). In addition, the replication will benefit from a better-established recognition programme. Three improvements will be introduced. Firstly, the hypotheses of the future iteration will be expanded to include predictions for the relevant subsamples, in order to offer a clear boundary between experimental hypotheses and possible exploratory activities. In like manner, attrition and outliers criteria have been made explicit (Nosek et al., 2018). Secondly a simple follow up protocol for measuring awareness has been approved by the institution. This measurement will be used to model the impacts of awareness on the treatment effect. Thirdly, the intervention will continue throughout the entire quarterly cycle, in order to capture the effects of late registry and execution of software tests.

The findings of this research also open new questions to improve the use of salience in different behavioural management devices. One first line of research could explore the differential effectiveness of salience when applied through alternative communication channels. For example, a recency device can be introduced in one the various scrum master meetings or applied through visual aids on digital channels such as emails or intranet sites. A second line of research could explore the effectiveness of recency on other behavioural management techniques such as feedback. Feedback is indispensable for orienting performance improvement, personal development and career planning (Cummings & Worley, 2014), but employees believe their productivity suffers due to lack of adequate feedback (Zenger & Folkman, 2014). Recency devices could be introduced in randomly assigned programmes in an attempt to stimulate the delivery of feedback.

As the application of behavioural effects become more prominent in recognition programmes, the researcher is also interested in the effects of loss aversion. The current recognition device starts every quarter with cero marks assigned to scrum masters, and they start accumulating marks as different activities are correctly completed. Because losses loom larger than gains (Kahneman & Tversky, 2000), the institutional scoring system could depart from 100 marks initially for every scrum master, and marks can be deducted as the expected activities are not completed. The cluster randomisation exercise discussed in the Methodology section could be used to measure this effect, and complete programmes could be assigned to one of two conditions: scoring with additive accumulation or with subtractive deduction.

7. Final Remarks

Recency mechanisms offer a promising means to trigger behavioural adoption, and there are still many fields where their use is yet to be scientifically explored. As of today, recency effects have found to operate in contexts such as taxation, political communications, and consumer behaviour, and recognition programmes offers a one novel field where recency could be applied.

The findings suggest that recency increased the effectiveness of the underlying recognition programme by 8% (p=0.05) among a broad pool of project teams, even after exploring possible covariates and controlling for their effect. A higher effect of 8.8% (p=0.035, also after controlling for covariates), was found among teams with size and work scopes that follow theoretical recommendations. These findings suggest that recency could offer a valuable complement for accelerating behavioural adoption when combined with recognition programmes. Some support is also offered to the notion that recency effects could be achieved through routine institutional digital communications that follow guidelines similar to those proposed here, and that recency effects could be effective in similar institutional contexts.

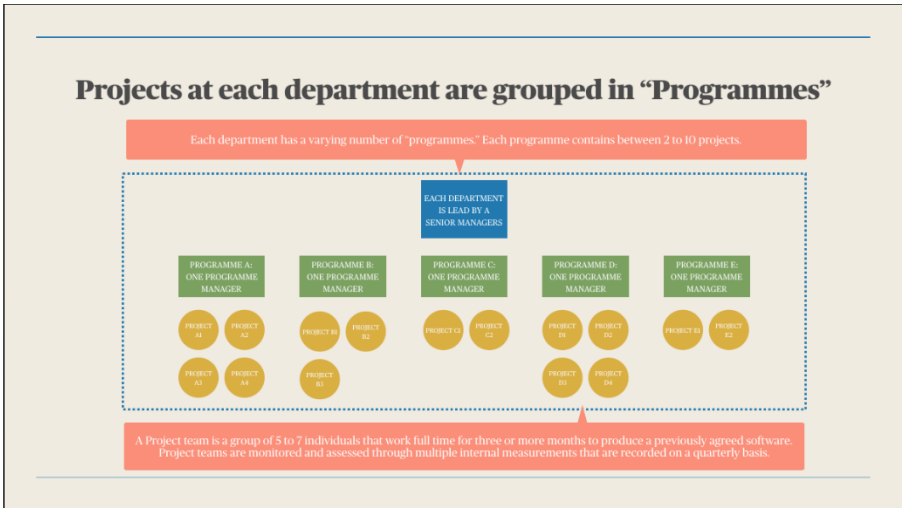
This research offers two practical implications in the policy making domain. Firstly, practitioners could take greater advantage of recency effects when deploying recognition programmes. Communicating recognition strategies relying solely on reflective mental processes could lead to lower success rates. The findings suggest that complementing such “changing minds” strategies with recency, which exploits the automatic and effortless nature of our minds can offer impactful results with a very low investment. Secondly, this research offers a reliable and innovative means to measuring behavioural change in institutional contexts, which is a current challenge for practitioners. In environments when multiple interventions are taking place, randomisation techniques can help discern an unbiased impact of behavioural management approaches. Practitioners in domains like marketing have extensively adopted this approach to test different strategies, and there is no reason why this should also be the case among professionals in other managerial domains such as human resources management.

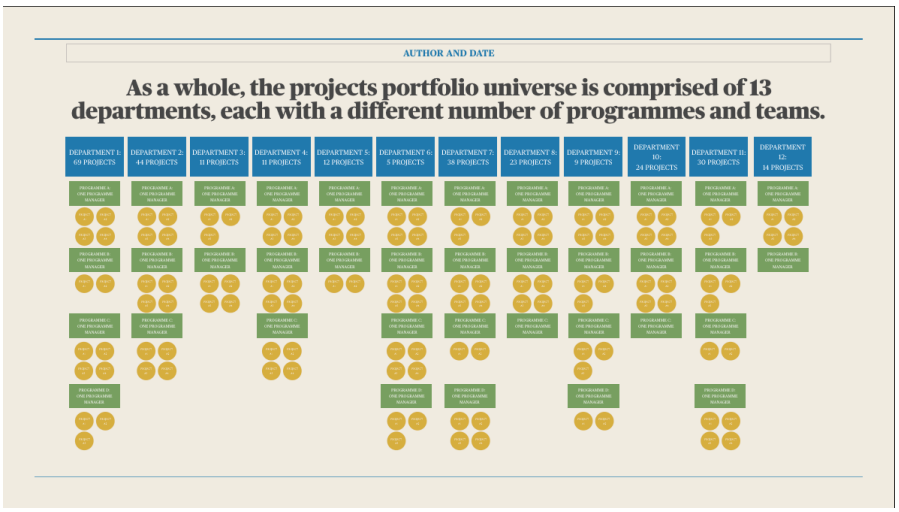
The encouraging findings of this intervention suggest further research avenues to understand other applications where recency can contribute to behavioural change. One avenue relates to the effectiveness of salience when applied through different communications channels, either individually or in parallel. This research explored the case of email communications, but work meetings, physical visual aids in the workplace, or other digital means could offer different average treatment effects. Another avenue could explore whether recency devices could be introduced to stimulate the delivery of other behavioural management techniques such as feedback, which is indispensable for performance but sadly insufficient in organisations.

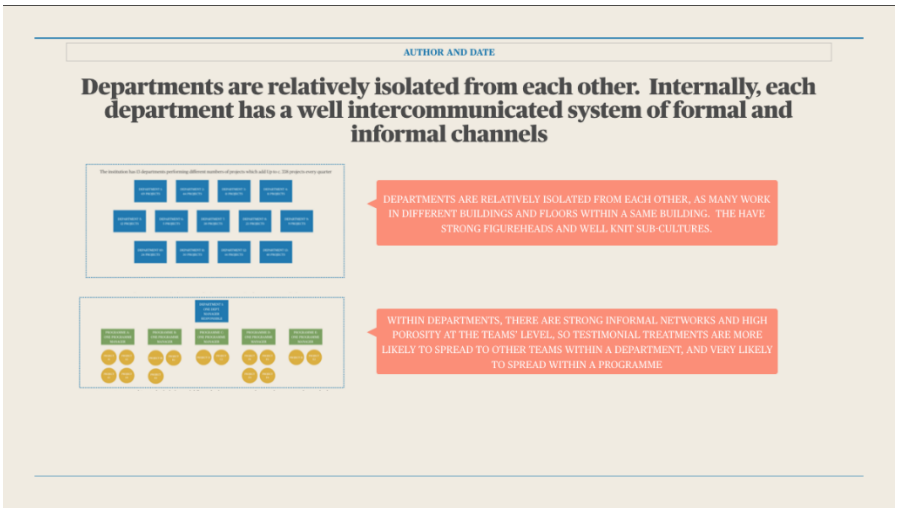
Behavioural Science, the science of what we do and how to change it, has revolutionized various domains of psychological and economic thought over the past 40 years. It has contributed to understand that modifying behaviours by “changing our minds” can be problematic in many ways. Conversely, Behavioural Science offers a wide array of effective mechanisms that rely on automatic mental processes that can be applied in a vast array of contexts where changing the way we behave will decisively improve our wellbeing. Hopefully, this research can make a humble contribution to such a noble objective.

8. Appendices

8.1. Appendix 1: Alternative experimental designs

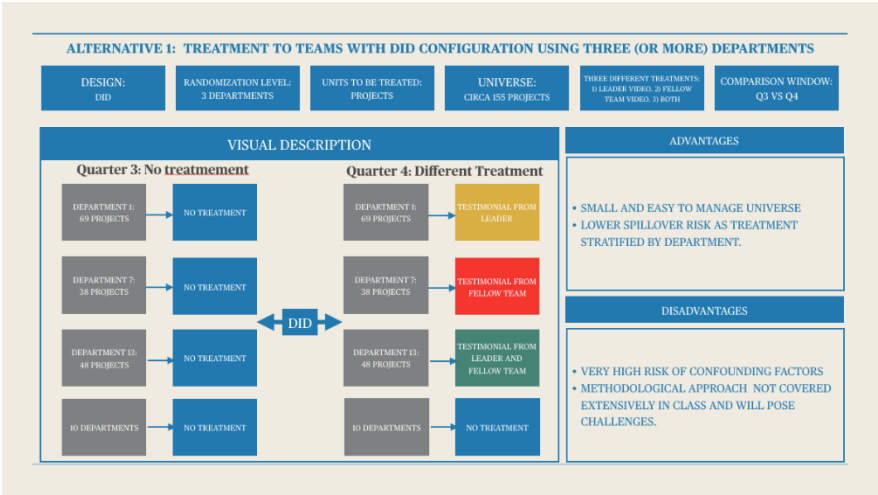

Projects at each department are grouped in "Programmes"
Each department has a varying number of "programmes." Each programme contains between 2 to 10 projects.
EACH DEPARTMENT IS LEAD BY A SENIOR MANAGERS
PROGRAMME A: ONE PROGRAMME MANAGER
PROGRAMME B: ONE PROGRAMME MANAGER
PROGRAMME C: ONE PROGRAMME MANAGER
PROGRAMME D: ONE PROGRAMME MANAGER
PROGRAMME E: ONE PROGRAMME MANAGER
A Project team is a group of 5 to 7 individuals that work full time for three or more months to produce a previously agreed software. Project teams are monitored and assessed through multiple internal measurements that are recorded on a quarterly basis.

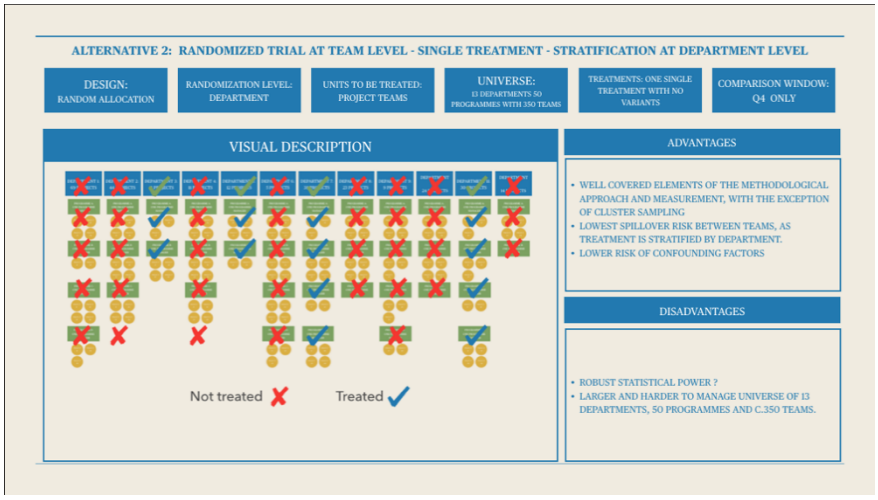

AUTHOR AND DATE
As a whole, the projects portfolio universe is comprised of 13 departments, each with a different number of programmes and teams.

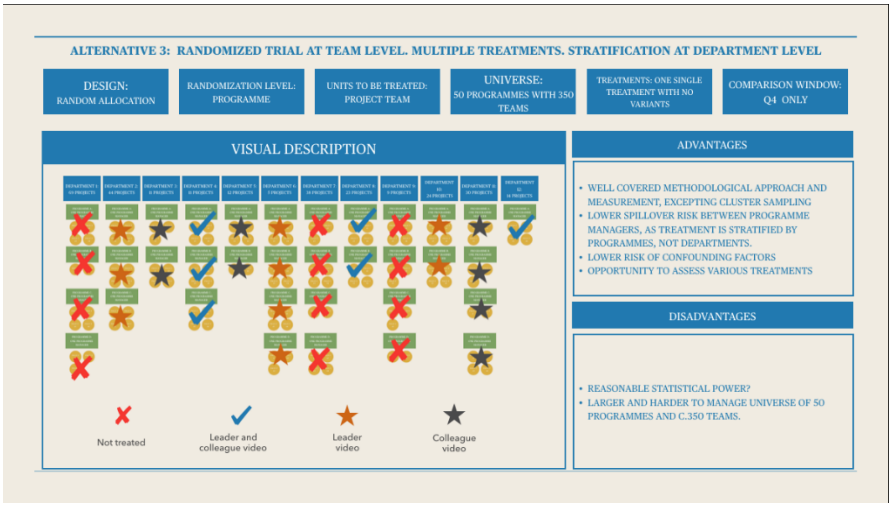

AUTHOR AND DATE
Departments are relatively isolated from each other. Internally, each department has a well intercommunicated system of formal and informal channels
DEPARTMENTS ARE RELATIVELY ISOLATED FROM EACH OTHER, AS MANY WORK IN DIFFERENT BUILDINGS AND FLOORS WITHIN A SAME BUILDING. THE HAVE STRONG FIGUREHEADS AND WELL KNIT SUB-CULTURES.
WITHIN DEPARTMENTS, THERE ARE STRONG INFORMAL NETWORKS AND HIGH POROSITY AT THE TEAMS' LEVEL, SO TESTIMONIAL TREATMENTS ARE MORE LIKELY TO SPREAD TO OTHER TEAMS WITHIN A DEPARTMENT, AND VERY LIKELY TO SPREAD WITHIN A PROGRAMME

Alternative experimental designs

Fact information






ALTERNATIVE 3: RANDOMIZED TRIAL AT TEAM LEVEL. MULTIPLE TREATMENTS. STRATIFICATION AT DEPARTMENT LEVEL
DESIGN: RANDOM ALLOCATION
RANDOMIZATION LEVEL: PROGRAMME
UNITS TO BE TREATED: PROJECT TEAM
UNIVERSE: 50 PROGRAMMES WITH 350 TEAMS
TREATMENTS: ONE SINGLE TREATMENT WITH NO VARIANTS
COMPARISON WINDOW: Q4 ONLY
VISUAL DESCRIPTION
Not treated
Leader and colleague video
Leader video
Colleague video
ADVANTAGES
• WELL COVERED METHODOLOGICAL APPROACH AND MEASUREMENT, EXCEPTING CLUSTER SAMPLING
• LOWER SPILLOVER RISK BETWEEN PROGRAMME MANAGERS, AS TREATMENT IS STRATIFIED BY PROGRAMMES, NOT DEPARTMENTS.
• LOWER RISK OF CONFOUNDING FACTORS
• OPPORTUNITY TO ASSESS VARIOUS TREATMENTS
DISADVANTAGES
• REASONABLE STATISTICAL POWER?
• LARGER AND HARDER TO MANAGE UNIVERSE OF 50 PROGRAMMES AND C.350 TEAMS.

8.2. Appendix 2: Detailed Stata outputs of cluster randomization scenarios

```
. clustersampsi, samplesize mu1(50) mu2(54) sd1(11) sd2(11) m(4) rho(0.9)
Sample size calculation determining the number of clusters required,
for a two sample comparison of means (using normal approximations).

For the user specified parameters:

mean 1:                                                 50.00
mean 2:                                                 54.00
standard deviation 1:                                   11.00
standard deviation 2:                                   11.00
significance level:                                     0.05
power:                                                  0.80
baseline measures adjustment (correlation):             0.00
average cluster size:                                   4
intra cluster correlation (ICC):                        0.9000
coefficient of variation (of cluster sizes):            0.00


clustersampsi estimated parameters:


Firstly, assuming individual randomisation:
sample size per arm:                                    119

Then, allowing for cluster randomisation:
design effect:                                          3.70
sample size per arm:                                    444
number clusters per arm:                                111

Note: sample size per arm required under cluster randomisation is rounded up
to a multiple of average cluster size and includes the addition
of one extra cluster per arm (to allow for t-distribution).
To understand sensitivity to these conservative allowances:
power with m clusters per arm:                          0.80
power with m-1 clusters per arm:                        0.80
```

```
. clustersampsi, samplesize mu1(50) mu2(54) sd1(6) sd2(6) m(4) rho(0.9)
Sample size calculation determining the number of clusters required,
for a two sample comparison of means (using normal approximations).

For the user specified parameters:

mean 1:                                                 50.00
mean 2:                                                 54.00
standard deviation 1:                                   6.00
standard deviation 2:                                   6.00
significance level:                                     0.05
power:                                                  0.80
baseline measures adjustment (correlation):             0.00
average cluster size:                                   4
intra cluster correlation (ICC):                        0.9000
coefficient of variation (of cluster sizes):            0.00


clustersampsi estimated parameters:


Firstly, assuming individual randomisation:
sample size per arm:                                    36

Then, allowing for cluster randomisation:
design effect:                                          3.70
sample size per arm:                                    136
number clusters per arm:                                34

Note: sample size per arm required under cluster randomisation is rounded up
to a multiple of average cluster size and includes the addition
of one extra cluster per arm (to allow for t-distribution).
To understand sensitivity to these conservative allowances:
power with m clusters per arm:                          0.80
power with m-1 clusters per arm:                        0.79
```

8.3. Appendix 3: Full Ethics Review Form Submitted to the Research Ethics Committee

This form should be completed for every research project that involves human participants or the use of information relating to directly identifiable individuals.

PART I - Checklist

The Checklist is designed to identify the nature of any ethical issues raised by the research.
This checklist must be completed before potential participants are approached to take part in any research.

1. Name of Researcher:

	Status *(mark with an 'X' as appropriate)*	Undergraduate student		Masters student	X
		Research degree student		Staff	
	Email	j.c.valverde-solano@lse.ac.uk	Telephone number	+525613688468	
	Department	Psychological and Behavioural Sciences			

2. Student Details if applicable

	Degree programme:	EMsC Behavioural Science Psychological and Behavioural Sciences		
	Supervisor's name:	Matteo M Galizzi	Supervisor's email:	m.m.galizzi@lse.ac.uk
	Supervisor's department:	Psychological and Behavioural Sciences		

3. Title of the proposal and brief abstract:

i) Title: Behavioural change through ego effects: Modifying Culture in institutional contexts.

ii) Abstract

Individuals behave in ways that make them feel good about themselves. They strive to maintain and enhance their self-esteem, and to achieve and maintain a positive social identity (Tajfel et al., 1979, p. 40). Such effects are referred to as "ego effects" in behavioural Science.

The primary objective of the Research project is to investigate the effect of ego-related psychological mechanisms in the adoption of novel work-related practices in institutional contexts. The dissertation will investigate if there is a difference in the adherence to novel work practices between software developers exposed to an emailed postcard motivating them to "be a top-performer" with regards to such novel practice, and those that are not exposed to such postcard.

I propose a between-subjects randomised controlled natural experiment using complete randomisation (Haynes et al., 2012). I will randomly select Treatment and Control samples of project teams that fulfil a set of minimum criteria. Individuals in the Treatment group will receive an email postcard motivating them to "be a top-performer" with regards to such novel practice, whilst Control Group individuals will not receive such postcard.

The effect of the treatment will be measured by the variation in the adoption of the novel cultural practice between Treatment and Control Groups.

4. Funding

Is it proposed that the research will be funded? No, no funding is necessary
If so by whom?

5. Where the research will be conducted

In what country/ies will the research take place? (See Note 1) Mexico City, Mexico
6. Data Management Plans
Please confirm whether you have completed a Data Management Plan and submitted to Datalibrary@lse.ac.uk ? (*See Note 2*) Yes / No

	Please mark an X in the appropriate right-hand column/box	**Yes**	**No**	**Not certain**
7. Research that *may* need to be reviewed by an external (non-LSE) Ethics Committee				
i	Will the study require Health Research Authority approval? *(See Note 3)*		X	
ii	Does the study involve participants lacking capacity to give informed consent? *(See Note 4)*		X	
iii	Is there any other reason why the study may need to be reviewed by another external (non-LSE) Ethics Committee? If yes, please give details here:		X	
If your research will be reviewed by an external (non-LSE) ethics committee, you may not need to complete the rest of this LSE review form – please email research.ethics@lse.ac.uk for guidance.				
8. Consent (See Note 5)				
i	Does the study involve children or other participants who are potentially or in any way vulnerable or who may have any difficulty giving meaningful consent to their participation or the use of their information? *(See Note 6)*		X	
ii	Are subjects to be involved in the study without their knowledge and consent (e.g. through internet-mediated research, or via covert observation of people in public places)?	X		
iiii	Will the study require the co-operation of a gatekeeper for initial access to the groups or individuals to be recruited? *(Answer 'yes' to this question only if the involvement of a gatekeeper in your study might raise issues of whether participants' involvement is truly voluntary or of whether the gatekeeper might influence potential participants in some other way.)*		X	
9. Research Design / Methodology				
i	Does the research methodology involve the use of deception? *(See Note 7)*			
ii	Are there any significant concerns regarding the design of the research project? For example: • where research intrudes into the private sphere or delves into some deeply personal experience; • where the study is concerned with deviance or social control; • where the study impinges on the vested interests of powerful persons or the exercise of coercion or domination; or • where the research deals with things that are sacred to those being studied that they do not wish profaned.		X	

	Please mark an X in the appropriate right-hand column/box	**Yes**	**No**	**Not certain**
iii	Does the proposed research relate to the provision of social or human services?		X	
10. Financial Incentives				
	Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants that might have an impact on the objectivity of the research?		X	
11. Research Subjects				
i	Could the study induce unacceptable psychological stress or anxiety or cause harm or negative consequences beyond the risks encountered in normal life?		X	
ii	Will the study involve discussion of sensitive topics? For example (but not limited to): sexual activity, illegal behaviour, experience of violence or abuse, drug use, etc.). (Please refer to the Research Ethics Policy, § 13).		X	
iii	Are drugs, placebos or other substances to be administered to study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?		X	
12. Confidentiality				
i	Will research involve the sharing of data or confidential information beyond the initial consent given?		X	
ii	Is there ambiguity about whether the information/data you are collecting is considered to be public?		X	
iii	Will the research involve administrative or secure data that requires permission from the appropriate authorities before use?		X	
iv	Will the research involve the use of visual/vocal methods that potentially pose an issue regarding confidentiality and anonymity?		X	
13. Legal requirements				
	Is there any reason why the research will NOT comply with the requirements of current data protection legislation? *(See Note 8)*		X	
14. Dissemination				
	Are there any particular groups who are likely to be harmed by dissemination of the results of this project? Or is there any potential for misuse of the findings?		X	
15. Risk to researchers				
	Does your research pose any risks to your physical or psychological wellbeing, or that of others working with you?		X	
16. Sensitive research materials				
	Will the research involve accessing security-sensitive material, such as material related to terrorism or violent extremism of any kind? (See Note 9)		X	

PART II: Low Risk, Departmental/centre/institute certification and/or next steps

Please note that there are certain circumstances where Departmental-certification of ethics review is not appropriate. Please see Note 10.

A If, after careful consideration, you have answered **No** to all the questions, you do not need to complete the questionnaire in Part III, unless you are subject to some external requirement that requires you to seek formal approval from the School's Research Ethics Committee. You can select **A** in the **Low risk, Departmental/Centre/Institute Certification Section** below, sign as appropriate and submit the form to the appropriate approver in your Department, Centre or Institute. Occasional audits of such forms may be undertaken by the School.

B If you have answered **Yes** or **Not certain** to any of the questions in sections 8-16 of the checklist you will need to consider more fully how you plan to deal with the ethical issues raised by your research. Answering the relevant questions in the Questionnaire in Part III below may assist you. If having done so you are wholly assured that adequate safeguards in relation to the ethical issues raised can and will be put in place, you may select **B** in the Low risk/Departmental/Centre/Institute certification Section below, sign as appropriate and submit the form to the appropriate approver in your Department, Centre or Institute. Occasional audits of such forms may be undertaken by the School.

C If you have answered Yes in section 7 that your research will be subject to review by an external (non-LSE) ethics committee, please select **C** below and send the Checklist (questions 1-7) to research.ethics@lse.ac.uk. You should submit your research for ethics approval to the appropriate external body. Once approval is granted please send a copy of the letter of approval to research.ethics@lse.ac.uk.

D If Departmental/Centre/Institute certification is not appropriate you should complete the questionnaire in Part III below, the '**Refer to Research Ethics Committee Section'** at the end of the form, and then submit the form to research.ethics@lse.ac.uk

LOW RISK, DEPARTMENTAL/CENTRE /INSTITUTE CERTIFICATION

Select A, B or C *(delete as appropriate):*

I have read and understood the LSE Research Ethics Policy and the questions contained in the Checklist above and confirm:

A that no significant ethical issues are raised by the research, or

B that adequate safeguards in relation to such issues can and will be put in place, or

C that the research will be subject to an external ethics review

Please complete the box below and sign the relevant section

i) Summary of any ethical issues identified and safeguards to be taken

ii) Details of relevant experience or training in this area

Low risk/Departmental / Centre / Institute Certifications should be approved as follows:

- MSc (or undergraduate) student review forms should be approved/signed by the academic supervisor. (PhD students cannot approve ethics review forms);
- PhD student review forms should be approved/signed by the supervisor
- Research staff who are not PIs should have their review forms approved/signed by the PI;

• Faculty and any research staff who are PIs on grants should have their review forms counter-signed by a designated research ethics champion in their Department / Centre or Institute, for example its research director

Signature of researcher (whether student or staff):		Date:	
Approved by (name)			
Approved by (signature)*:		Date:	

*By signing here the approver confirms that to the best of their understanding any ethical issues have been adequately addressed in the research design, and the researcher has been made aware of her/his responsibilities for the ethical conduct of her/his research. If in doubt, please refer to your departmental ethics champion, or to the Research Governance Manager, research.ethics@lse.ac.uk

Part III - Questionnaire

The questionnaire enables you to explain how the ethical issues relating to your research will be addressed. If you are intending to submit your proposal to the Research Ethics Committee it needs to be completed in full.

17. Research aims

Please provide brief (no more than approx.500 words) details in non-technical language of the research aims, the scientific background of the research and the methods that will be used. This summary should contain sufficient information to acquaint the Committee with the principal features of the proposal. A copy of the full proposal should nonetheless be attached to this document in case it is required for further information.

Institution in different contexts are struggling to adopt cultural changes in a faster way. At the same time, most leadership groups lack a robust understanding of the proven ways to address such challenges (Dichter et al., 2019). This research can contribute to this debate by helping to elucidating the effectiveness of a behavioural change technique within the context of cultural change.

Ego effects can be used to effect change individual behaviour within institutional contexts. Individuals behave in ways that make them feel good about themselves. They strive to maintain and enhance their self esteem, and to achieve and maintain a positive social identity (Tajfel et al., 1979, p. 40). In a famous example of this, male donors have been found to donate more to charity when approached by attractive females, as a result of desiring to maintain a positive self-image in the eyes of the opposite sex (Landry et al., 2006). Individuals' pursuit of a positive self-image also leads them to compare against others and 'self-evaluate', and to calibrate their behaviour (Tesser, 1986). This process can be triggered through bespoke feedback that can effectively modify behaviour (Harper, 2019).

I propose an experiment in a natural context where software development teams[1] code and implement new Software in a large private institution. This organisation is attempting to embed Agile software development practices among a pool of approximately 300 Scrum Teams. The experiment will take advantage of the firm's need to adopt the practice of performing a new project planning activity, which is not part of the current culture.

The primary objective of the Research project is to investigate the effect of ego-related psychological mechanisms in the adoption of novel work-related practices in institutional contexts. Is there is a difference in the adherence to novel work practices between software developers exposed to an emailed postcard motivating them to "be a top-performer" with regards to such novel practice, and those that are not exposed to such postcard?

I will use a between-subjects randomised controlled natural experiment using complete randomisation (Haynes et al., 2012). I will randomly select Treatment and Control samples of project teams that fulfil a set of minimum criteria. Individuals in the Treatment group

[1] The terms "project," "team," "project team," and "Scrum," will be used interchangeably to refer to a Scrum team: a group of 4 to 10 individuals that work exclusively and full-time, for one or more quarters to produce a previously agreed software product.

will receive an email postcard motivating them to "be a top-performer" with regards to such novel practice, whilst Control Group individuals will not receive such postcard. The research is relevant because a preliminary review of the literature did not reveal an application of ego effects in the introduction of new cultural practices in institutional contexts. This experiment will enhance the body of knowledge with regard to the ego-related psychological mechanisms in the adoption of novel work-related practices.	
18. Informed consent	
i.	Has information (written and oral) about the study been prepared in an appropriate form and language for potential participants? At what point in the study will this information be offered? (See Note 5)
	Subjects will be involved in the study without their knowledge and consent.
ii	Will potential participants be asked to give informed consent ***in writing*** and will they be asked to confirm that they have received and read the information about the study? If not, why not? *Please attach your proposed information sheet and consent form.*
	Subjects will be involved in the study without their knowledge and consent.
iii.	If the research takes place within an online community, explain how informed consent will be obtained? What arrangements are in place for ensuring that participants do not include vulnerable groups or children?
	Subjects will be involved in the study without their knowledge and consent.
iv.	How has the study been discussed or are there plans to discuss the study with those likely to be involved, including potential participants or those who may represent their views?
	Subjects will be involved in the study without their knowledge and consent.
v	Will potential participants be clearly informed that no adverse consequences will follow a decision not to participate or to withdraw during the study?
	Subjects will be involved in the study without their knowledge and consent.
vi	What provision has been made to respond to queries and problems raised by participants during the course of the study?
	Subjects will be involved in the study without their knowledge and consent.
19. Research design and methodology	
i	Where the research involves the use of deception (or the withholding of full information about the study), how does the research methodology justify this?
	No deception is involved.
ii	How will data be collected and analysed during the project?
	Data will be collected by the researcher from the institutional management information systems. I propose the following analyses to estimate the average treatment effect: • Basic descriptive statistics and graphical description of dependent variable and other variables. • Pairwise correlation between the dependent variables and other variables. • Normality test of the dependent variable using a Shapiro-Wilk test. • Statistical tests of $\mathbf{H_0}$, either with t-test or Mann-Whitney according to normality results. • Linear simple and multiple regressions with an OLS model, to test $\mathbf{H_0}$ and to determine the effect of the treatment for the entire sample and relevant subsamples.

	• Ex-post sample size calculation to detect the treatment effect.
iii	How have the ethical and legal dimensions of the process of collecting, analyzing and storing the data been addressed?
	The project has been discussed with Senior management at the firm where it will take place and they have agreed to it.
iv	If agencies, communities or individuals may be directly affected by the research (e.g. participants, service users, vulnerable communities or relations), what means have you devised to ensure that any harm or distress is minimized and/or that the research is sensitive to the particular needs and perspectives of those so affected?
	No possible harm to communities is identified at this moment.
20. Ethical questions arising from the provision of incentives	
	Are any incentives being offered to participants? If so, please provide details
	No incentives are offered.
21. Research participants	
i	Who do you identify as the participants in the project? Are other people who are not participants likely to be directly or indirectly impacted by the project?
	Participants are not directly impacted by the project. The participants are a random sample of employees at the firm. Participants work in the software development department of the firm.
ii	Are there any specific risks to research participants or third parties? If so, please give details
	No specific risks have been identified.
iii	If the research involves pain, stress, physical or emotional risk, please detail the steps taken to minimize such effects.
	No pain, stress, physical or emotional risk is involved.
22. Confidentiality	
i.	What arrangements have been made to preserve confidentiality and anonymity for the participants or those potentially affected, and compliance with data protection law?
	All data will be anonymous. Names, locations, groups, project names, will be confidential and not revealed by the researcher.
ii	Have you considered the limits to confidentiality, if, for instance, a participant should disclose information which suggests that they or someone else may be at significant risk of harm?
	Such situations are not considered in the study.
23. Dissemination	
	Will the results of the study be offered to those participants or other affected parties who wish to receive them? If so, what steps have been taken to minimize any discomfort or misrepresentation that may result at the dissemination stage.
	The results of the study will only be shared with senior management at the firm where it will take place.
24. Risk to researchers	
	Are there any risks to researchers? If so, please provide details.
	No risks to researchers are identified at the moment.

REFER TO RESEARCH ETHICS COMMITTEE		
Approval is required by the Research Ethics Committee on one or more of the following grounds (please mark with an 'X' in the appropriate place in the right-hand column):		
a.	Significant ethical issues are raised by the research, including research characterised by one or more of the following features: (i) Research involving deception of participants, or which is conducted without their full and informed consent at the time the study is carried out or when the data is gathered, or which involves the use of confidential information. (ii) Research where informed consent will be obtained orally but not in writing; (iii) Research involving any of the following: vulnerable groups; personally intrusive or ethically sensitive topics; groups where permission of a gatekeeper is normally required for initial access to members (where involvement of the gatekeeper might raise issues of whether participants' involvement is truly voluntary); research which would induce undue psychological stress, anxiety or humiliation cause more than minimal pain; (iv) Research involving more than minimal risk of harm to the researcher(s)	
b.	The researcher wants to seek the advice of the Research Ethics Committee	
c.	External obligations (for instance, funder requirements, data access requirements) require it	
d.	Research undertaken by a student or member of staff who has not received appropriate training or has insufficient experience in research ethics and has been unable to access appropriate advice or support.	
Please submit your review form, research proposal and your planned Information Sheet and Consent form to research.ethics@lse.ac.uk for review by the Research Ethics Committee.		

NOTES

1. If the research will be conducted abroad you will need to complete a Notification to Travel form. If you will be travelling to a high risk destination you may need to complete a risk identification form and a risk assessment form. Please see: https://info.lse.ac.uk/staff/divisions/Risk-and-Compliance-Unit/Health-and-Safety/Fieldwork-overseas-travel-and-off-site-activities . Note that if the location or nature of the research presents a high degree of risk, the Research Ethics Committee may check with the Health and Safety team that a risk assessment is underway.

2. If you have not already done so, please complete a Data Management Plan (DMP). We recommend using the templates provided on DMPonline: https://dmponline.dcc.ac.uk/ Guidance on writing a DMP and using DMPonline can be found on the Library webpages at:
http://www.lse.ac.uk/Library/Research-support/Research-Data-Management/What-is-a-Data-Management-Plan-and-how-do-I-write-one Unless you have a research funder that is listed, selected the generic DMP option. Please submit your completed DMPs to the Data Librarian on Datalibrary@lse.ac.uk

3. If your research involves participants identified from, or because of, their status as patients of the NHS or other health services of the UK Devolved Administrations, and/or the relatives of such patients then it will most likely fall under the remit of the Health Research Authority; similarly, social care research involving adults children or families and some proposals for social science studies situated in the NHS will fall under the remit of the Social Care Research Ethics Committee. There is an easy-to-use tool to help you ascertain whether or not you need HRA approval or not at: http://www.hra-decisiontools.org.uk/ethics/ For further guidance see: http://www.hra.nhs.uk/research-community/before-you-apply/determine-which-review-body-approvals-are-required/

4. Under the Mental Capacity Act 2005, research involving adults aged 16 or over with learning difficulties or who otherwise 'lack capacity' will be subject to approval by an NHS REC if that research is deemed to be 'intrusive'. For guidance see: http://www.hra.nhs.uk/resources/research-legislation-and-governance/questions-and-answers-mental-capacity-act-2005/

5. Please refer to the LSE guidance on Informed Consent (which includes a sample template) here:
http://www.lse.ac.uk/intranet/LSEServices/policies/pdfs/school/infCon.pdf. Note that if you will **not** be obtaining **written** consent then your ethics application will need to be submitted to the Research Ethics Committee for review.

6. Please note that we follow the ESRC definition of vulnerability which is as follows: 'Vulnerability may be defined in different ways and may arise as a result of being in an abusive relationship, vulnerability due to age, potential marginalisation, disability, and due to disadvantageous power relationships within personal and professional roles. Participants may not be conventionally 'vulnerable', but may be in a dependent relationship that means they can feel coerced or pressured into taking part, so extra care is needed to ensure their participation is truly voluntary.' https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/frequently-raised-topics/research-with-potentially-vulnerable-people/
Please also note that as general guidance, research participants under the age of 18 may be vulnerable. If your research will involve children or other potentially vulnerable participants please refer to the LSE Safeguarding policy at:
https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/safPol.pdf
Also, see Note 4 above regarding the Mental Capacity Act.

7. Deception can occur at a variety of levels: for example, at one level, experimental methods may depend on participants being deliberately misled as to the true nature or purpose of the research in which they are taking part; at another, covert participant observation may entail an implicit deception as to the true identity and role of the researcher. Deception may be a legitimate and necessary feature of social scientific research, but its use must always be properly justified. Any research involving deception must be submitted to the LSE Research Ethics Committee for review.

8. Please refer to the School's guidance on Data Protection and research: https://info.lse.ac.uk/staff/services/Policies-and-procedures/Assets/Documents/datProRes.pdf

9. Where staff or students are planning research projects that will entail accessing security-sensitive material, it is important we ensure that the necessary safeguards are in place to protect both the researcher and the School. Even where there are no ethical issues raised by the research (inasmuch that there are no human participants) it is very important that we have a log of any such research so that students or staff do not run the risk of being wrongly accused of accessing such materials for other/non-research reasons. If your research will involve accessing such material please email research.ethics@lse.ac.uk

10. Applications relating to the following kinds of research should always be subject to review by the Research Ethics Committee:
(i) Research involving deception of participants, or that is intentionally conducted without their full and informed consent at the time the study is carried out or when the data are gathered
(ii) Research which involves or may lead to the publication of confidential information
(iii) Research where informed consent will be obtained orally but not in writing
(iv) Research involving any of the following:
• research involving vulnerable groups ;
• research involving sensitive topics ;
• research involving groups where permission of a gatekeeper is normally required for initial access to members (where involvement of the gatekeeper might raise issues of whether participants' involvement is truly voluntary);
• research which would induce undue psychological stress, anxiety or humiliation or cause more than minimal pain.
(v) Research involving more than minimal risk of harm (whether emotional or physical) to the researcher(s)

8.4. Appendix 4: Approval from the Research Ethics Committee.

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Houghton Street
London WC2A 2AE
United Kingdom

tel: +44 (0)20 7852 3629
email: research.ethics@lse.ac.uk

www.lse.ac.uk

Research Ethics Committee

Juan Carlos Valverde
Department of Psychological and Behavioural Science
j.c.valverde-solano@lse.ac.uk

18th September 2020

Dear Juan Carlos

Re: 'Behavioural change through ego effects: Modifying Culture in institutional contexts'

REC ref. 1213

I am writing with reference to the above research proposal. The Research Ethics Committee, having considered the documentation sent, is satisfied that the ethical issues raised by the proposed research have been properly taken into account and that adequate safeguards have been put in place. I am accordingly able on behalf of the Committee to confirm our approval of the application.

Please note that any significant changes to the research design must be reported to the Research Ethics Committee. Amendments to the research design that may affect participants and/or that may have ethical implications must be reviewed and approved by the Research Ethics Committee before commencement (or recommencement) of the project. The Research Ethics Committee may periodically conduct a selective audit of current research projects.

I would like to take this opportunity to wish you well with your research project.
If you have any further queries please contact Lyn Grove via research.ethics@lse.ac.uk.

Yours sincerely,

David J. Lewis

Professor David Lewis
Chair, Research Ethics Committee
cc. Dr Lyn Grove, Research & Innovation Division

8.5. Appendix 5: Stata-Do File Used to Analyse of the Data

```
/* 18080 DISSERTATION DO FILE */
/* DATABASE COLLECTED ON 20 NOVEMBER 2020*/

/* Log File*/

. cd "/Users/jcvalverde/Dropbox/LSE Behavioural Science/15-PB451E Dissertation/3-Final Intervention/6) Do and txt Files/"

log using "18080 Dissertation FINAL .txt", text append

/* DISTRIBUTION OF TREATMENT AND CONTROL GROUPS */

tabulate Treatment

/* ATTRITION ANALYSIS */

tabulate Treatment Attrition

tabulate Treatment if Attrition!=1

/* DESCRIPTIVE STATISTICS OF THE DEPENDENT VARIABLE */

summarize  PercentageWithTest if Attrition!=1, detail

sort Treatment
by Treatment: summarize PercentageWithTest if Attrition!=1, detail

/* DESCRIPTIVE STATISTICS OF THE DEPENDENT VARIABLE IN THE CONTROL AND TREATMENT GROUPS */

sort Treatment
by Treatment: summarize PercentageWithTest if Attrition!=1, detail

/* GRAPHICAL REPRESENTATION OF THE DEPENDENT VARIABLE IN THE TREATMENT AND CONTROL GROUPS*/

histogram PercentageWithTest if Attrition!=1, width(10) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features with a software test task) title("Distributon of Scrums according to the percentage of Features with a software test task", span size(small)) subtitle( ”Salience of recognition strategies at the workplace”, span size(small)) note(As of the sixth week of the Quarter)  xlabel(0(10)100)

histogram PercentageWithTest if Treatment==0 & Attrition!=1, width(10) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features with a software test task) title("Control Group: Distributon of Scrums according to the percentage of Features with a software test task", span size(small)) subtitle( ”Subtitle pending”, span size(small)) note(As of the sixth week of the Quarter)  xlabel(0(10)100)

histogram PercentageWithTest if Treatment==1 & Attrition!=1, width(10) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features with a software test task) title("Treatment Group: Distributon of Scrums according to the percentage of Features with a software test task", span size(small)) subtitle( ”Subtitle pending”, span size(small)) note(As of the sixth week of the Quarter)  xlabel(0(10)100)

/* 5) PAIRWISE CORRELATION */

pwcorr PercentageWithTest Rank NewDummy TrainingDummy TotalFeatures TeamSize NewvsOngoingSMDummy if Attrition!=1, sig star (0.05)

/* NORMALITY TEST OF THE DEPENDENT VARIABLE*/

swilk PercentageWithTest if Attrition!=1

/* PARAMETRIC STATISTICAL TEST OF NULL HYPOTHESIS*/

ttest PercentageWithTest if Attrition!=1, by(Treatment)

/* GENERAL LINEAR REGRESSION */

regress PercentageWithTest Treatment if Attrition !=1

/* GENERAL MULTIPLE LINEAR REGRESSION */

regress PercentageWithTest Treatment TotalFeatures if Attrition !=1

/* TESTS OF ROBUSTNESS AMONG DIFFERENT SUBSAMPLES*/
```

```
sort NewDummy
by NewDummy:  ttest PercentageWithTest if Attrition!=1, by(Treatment)

sort NewvsOngoingSMDummy
by NewvsOngoingSMDummy: ttest PercentageWithTest if Attrition!=1, by(Treatment)

sort CompliantTeamSize
by CompliantTeamSize: ttest PercentageWithTest if Attrition!=1, by(Treatment)

sort CompliantProjectScope
by CompliantProjectScope: ttest PercentageWithTest if Attrition!=1, by(Treatment)

/*OUTLIERS ANALYSIS */

histogram TeamSize if Attrition!=1, width(1) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentage of total Teams ) ylabel
(,angle(horizontal)) xtitle(Number of Team Members) title("Distributon of Scrums according to Team Size", span size(small)) subtitle(
”Salience in Recogntion Programmes”, span size(small)) note(As of the sixth week of the Quarter)  xlabel(0(1)20)

histogram TotalFeatures if Attrition!=1, width(1) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentage of total Teams ) ylabel
(,angle(horizontal)) xtitle(Number of Features per Team) title("Distributon of Scrums according to Total Features", span size(small)) subtitle(
”Salience in Recogntion Programmes”, span size(small)) note(As of the 20th of November 2020)  xlabel(0(1)20)

/*PARAMETRIC TEST OF THE NULL HYPOTHESIS EXCLUDING OUTLIERS */

ttest PercentageWithTest if Attrition!=1 & TeamSize<=11 & TotalFeatures<=13, by(Treatment)

/* GENERAL LINEAR REGRESSION EXCLUDING OUTLIERS*/

regress PercentageWithTest Treatment if Attrition !=1 & TeamSize<=11 & TotalFeatures<=13

/* GENERAL MULTIPLE LINEAR REGRESSION EXCLUDING OUTLIERS*/

regress PercentageWithTest Treatment TotalFeatures if Attrition !=1 & TeamSize<=11 & TotalFeatures<=13

/* A POSTERIORI SAMPLE SIZE ESTIMATION */

sampsi 54.13 62.49, sd1(31.04) sd2(29.66) alpha(0.05) power(0.8)

/* TESTS TO VERIFY SAMPLES ARE NOT DIFFERENT AFTER ATTRITION*/

swilk Rank
ranksum Rank if Attrition!=1, by(Treatment)
return list

swilk NewDummy
tab NewDummy Treatment if Attrition!=1, chi2

swilk Q3Score
ranksum Q3Score if Attrition!=1, by(Treatment)
return list

swilk TotalFeatures
ranksum TotalFeatures if Attrition!=1, by(Treatment)
return list

swilk TeamSize
ranksum TeamSize if Attrition!=1, by(Treatment)
return list

swilk NewvsOngoingSMDummy
tab NewvsOngoingSMDummy Treatment if Attrition!=1, chi2

swilk TrainingDummy
tab TrainingDummy Treatment if Attrition!=1, chi2


log close

/* END OF DO FILE */
```

8.6. Appendix 6: Data output file with detailed calculations

```
-------------------------------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  /Users/jcvalverde/Dropbox/LSE Behavioural Science/15-PB451E Dissertation/3-Final Intervention/
> 6) Do and txt Files/18080 Dissertation FINAL .txt
  log type:  text
 opened on:  13 Dec 2020, 10:07:42

. 
. /* DISTRIBUTION OF TREATMENT AND CONTROL GROUPS */
. 
. tabulate Treatment

      Treatment |      Freq.     Percent        Cum.
----------------+-----------------------------------
  Control Group |        124       53.45       53.45
Treatment Group |        108       46.55      100.00
----------------+-----------------------------------
          Total |        232      100.00

. 
. /* ATTRITION ANALYSIS */
. 
. tabulate Treatment Attrition

                |       Attrition
      Treatment |         0          1 |     Total
----------------+----------------------+----------
  Control Group |       116          8 |       124
Treatment Group |        95         13 |       108
----------------+----------------------+----------
          Total |       211         21 |       232

. 
. tabulate Treatment if Attrition!=1

      Treatment |      Freq.     Percent        Cum.
----------------+-----------------------------------
  Control Group |        116       54.98       54.98
Treatment Group |         95       45.02      100.00
----------------+-----------------------------------
          Total |        211      100.00

. 
. /* DESCRIPTIVE STATISTICS OF THE DEPENDENT VARIABLE */
. 
. summarize  PercentageWithTest if Attrition!=1, detail

                    PercentageWithTest
-------------------------------------------------------------
      Percentiles      Smallest
 1%           0              0
 5%           0              0
10%       16.67              0       Obs                 211
25%       33.33              0       Sum of Wgt.         211

50%          60                      Mean           57.90256
                        Largest       Std. Dev.      30.63553
75%          80            100
90%         100            100       Variance        938.536
95%         100            100       Skewness      -.1626984
99%         100            100       Kurtosis       1.994017

. 
. sort Treatment

. by Treatment: summarize PercentageWithTest if Attrition!=1, detail

-------------------------------------------------------------------------------------------------------------------
-> Treatment = Control Group

                    PercentageWithTest
-------------------------------------------------------------
      Percentiles      Smallest
 1%           0              0
 5%           0              0
10%        12.5              0       Obs                 116
25%       30.95              0       Sum of Wgt.         116

50%          50                      Mean             54.1375
                        Largest       Std. Dev.      31.03528
75%          75            100
90%         100            100       Variance       963.1884
95%         100            100       Skewness       -.066121
```

```
99%          100            100       Kurtosis       1.985825

-------------------------------------------------------------------------------------------------------------------
-> Treatment = Treatment Group

                    PercentageWithTest
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0             0
 5%        14.29             0
10%           25             0       Obs                  95
25%        33.33         11.11       Sum of Wgt.          95

50%        66.67                     Mean           62.49989
                        Largest       Std. Dev.      29.65648
75%          100            100
90%          100            100       Variance       879.5067
95%          100            100       Skewness      -.2648247
99%          100            100       Kurtosis       2.023169


.
. /* DESCRIPTIVE STATISTICS OF THE DEPENDENT VARIABLE IN THE CONTROL AND TREATMENT GROUPS */
.
. sort Treatment

. by Treatment: summarize PercentageWithTest if Attrition!=1, detail

-------------------------------------------------------------------------------------------------------------------
-> Treatment = Control Group

                    PercentageWithTest
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0             0
 5%            0             0
10%         12.5             0       Obs                 116
25%        30.95             0       Sum of Wgt.         116

50%           50                     Mean            54.1375
                        Largest       Std. Dev.      31.03528
75%           75            100
90%          100            100       Variance       963.1884
95%          100            100       Skewness       -.066121
99%          100            100       Kurtosis       1.985825

-------------------------------------------------------------------------------------------------------------------
-> Treatment = Treatment Group

                    PercentageWithTest
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0             0
 5%        14.29             0
10%           25             0       Obs                  95
25%        33.33         11.11       Sum of Wgt.          95

50%        66.67                     Mean           62.49989
                        Largest       Std. Dev.      29.65648
75%          100            100
90%          100            100       Variance       879.5067
95%          100            100       Skewness      -.2648247
99%          100            100       Kurtosis       2.023169


.
. /* GRAPHICAL REPRESENTATION OF THE DEPENDENT VARIABLE IN THE TREATMENT AND CONTROL GROUPS*/
.
. histogram PercentageWithTest if Attrition!=1, width(10) start(0) fraction normal fcolor(red) gap(8) ytitl
> e(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features with a software
> test task) title("Distributon of Scrums according to the percentage of Features with a software test task
> ", span size(small)) subtitle( ,ÄùSalience of recognition strategies at the workplace,Äù, span size(small)) n
> ote(As of the sixth week of the Quarter)  xlabel(0(10)100)
(bin=10, start=0, width=10)

.
. histogram PercentageWithTest if Treatment==0 & Attrition!=1, width(10) start(0) fraction normal fcolor(re
> d) gap(8) ytitle(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features w
> ith a software test task) title("Control Group: Distributon of Scrums according to the percentage of Feat
> ures with a software test task", span size(small)) subtitle( ,ÄùSubtitle pending,Äù, span size(small)) note(A
> s of the sixth week of the Quarter)  xlabel(0(10)100)
(bin=10, start=0, width=10)

.
. histogram PercentageWithTest if Treatment==1 & Attrition!=1, width(10) start(0) fraction normal fcolor(re
> d) gap(8) ytitle(Percentage of total Scrums ) ylabel (,angle(horizontal)) xtitle(Percentage of Features w
> ith a software test task) title("Treatment Group: Distributon of Scrums according to the percentage of Fe
> atures with a software test task", span size(small)) subtitle( ,ÄùSubtitle pending,Äù, span size(small)) note
```

```
> (As of the sixth week of the Quarter)  xlabel(0(10)100)
(bin=10, start=0, width=10)

.
. /* 5) PAIRWISE CORRELATION */
.
. pwcorr PercentageWithTest Rank NewDummy TrainingDummy TotalFeatures TeamSize NewvsOngoingSMDummy if Attri
> tion!=1, sig star (0.05)

             | Percen~t     Rank NewDummy Traini~y TotalF~s TeamSize NewvsO~y
-------------+------------------------------------------------------------------
Percentage~t |   1.0000
             |
             |
        Rank |  -0.1196   1.0000
             |   0.0832
             |
    NewDummy |  -0.0144   0.1450*  1.0000
             |   0.8357   0.0353
             |
TrainingDu~y |  -0.0014  -0.0088  -0.1983*  1.0000
             |   0.9844   0.8993   0.0038
             |
TotalFeatu~s |  -0.2598* -0.1802* -0.1373*  0.0758   1.0000
             |   0.0001   0.0087   0.0464   0.2731
             |
    TeamSize |   0.1078  -0.4430* -0.1562*  0.0835   0.3090*  1.0000
             |   0.1185   0.0000   0.0233   0.2270   0.0000
             |
NewvsOngoi~y |   0.0159  -0.1576* -0.9289*  0.1447*  0.0829   0.1026   1.0000
             |   0.8186   0.0220   0.0000   0.0357   0.2302   0.1373
             |

.
. /* NORMALITY TEST OF THE DEPENDENT VARIABLE*/
.
. swilk PercentageWithTest if Attrition!=1

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+-------------------------------------------------------
Percentage~t |        211    0.99501      0.780    -0.572    0.71646

.
. /* PARAMETRIC STATISTICAL TEST OF NULL HYPOTHESIS*/
.
. ttest PercentageWithTest if Attrition!=1, by(Treatment)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |     116     54.1375    2.881553    31.03528     48.4297     59.8453
Treatmen |      95    62.49989    3.042691    29.65648    56.45856    68.54123
---------+--------------------------------------------------------------------
combined |     211    57.90256    2.109037    30.63553    53.74496    62.06016
---------+--------------------------------------------------------------------
    diff |           -8.362395    4.209696               -16.6613   -.0634859
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.9865
Ho: diff = 0                                     degrees of freedom =      209

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0241         Pr(|T| > |t|) = 0.0483          Pr(T > t) = 0.9759

.
. /* GENERAL LINEAR REGRESSION */
.
. regress PercentageWithTest Treatment if Attrition !=1

      Source |       SS           df       MS      Number of obs   =       211
-------------+----------------------------------   F(1, 209)       =      3.95
       Model |  3652.24974         1  3652.24974   Prob > F        =    0.0483
    Residual |    193440.3       209  925.551675   R-squared       =    0.0185
-------------+----------------------------------   Adj R-squared   =    0.0138
       Total |   197092.55       210  938.535952   Root MSE        =    30.423

------------------------------------------------------------------------------
Percentage~t |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   Treatment |   8.362395   4.209696     1.99   0.048     .0634859     16.6613
       _cons |    54.1375   2.824694    19.17   0.000     48.56896    59.70604
------------------------------------------------------------------------------

.
. /* GENERAL MULTIPLE LINEAR REGRESSION */
.
```

```
. regress PercentageWithTest Treatment TotalFeatures if Attrition !=1

      Source |       SS           df       MS      Number of obs   =       211
-------------+----------------------------------   F(2, 208)       =      9.61
       Model |  16670.4465         2  8335.22325   Prob > F        =    0.0001
    Residual |  180422.103       208  867.413958   R-squared       =    0.0846
-------------+----------------------------------   Adj R-squared   =    0.0758
       Total |   197092.55       210  938.535952   Root MSE        =    29.452

------------------------------------------------------------------------------
PercentageW~t |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
    Treatment |   8.037087   4.076203     1.97   0.050     .0011189    16.07305
TotalFeatures |  -3.538887   .9134916    -3.87   0.000    -5.339776   -1.737999
        _cons |   72.04549   5.370845    13.41   0.000     61.45722    82.63376
------------------------------------------------------------------------------

.
. /* TESTS OF ROBUSTNESS AMONG DIFFERENT SUBSAMPLES*/
.
. sort NewDummy

. by NewDummy:  ttest PercentageWithTest if Attrition!=1, by(Treatment)

-------------------------------------------------------------------------------------------------------------------
-> NewDummy = Ongoing Projects

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |      95    55.25389     3.19338    31.12521    48.91336    61.59442
Treatmen |      82    61.38634    3.266028    29.57514    54.88797    67.88471
---------+--------------------------------------------------------------------
combined |     177    58.09492    2.291443    30.48567    53.57267    62.61716
---------+--------------------------------------------------------------------
    diff |           -6.132447    4.585033               -15.18152    2.916631
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.3375
Ho: diff = 0                                     degrees of freedom =      175

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0914         Pr(|T| > |t|) = 0.1828          Pr(T > t) = 0.9086

-------------------------------------------------------------------------------------------------------------------
-> NewDummy = New Projects

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |      21    49.08714    6.734315    30.86051    35.03961    63.13468
Treatmen |      13    69.52385    8.427971    30.38748    51.16087    87.88682
---------+--------------------------------------------------------------------
combined |      34    56.90118    5.462705    31.85277    45.78722    68.01513
---------+--------------------------------------------------------------------
    diff |            -20.4367    10.82854               -42.49371    1.620305
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.8873
Ho: diff = 0                                     degrees of freedom =       32

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0341         Pr(|T| > |t|) = 0.0682          Pr(T > t) = 0.9659

.
. sort NewvsOngoingSMDummy

. by NewvsOngoingSMDummy: ttest PercentageWithTest if Attrition!=1, by(Treatment)

-------------------------------------------------------------------------------------------------------------------
-> NewvsOngoingSMDummy = New Scrum Masters

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |      18    49.86111    7.835457    33.24303    33.32974    66.39248
Treatmen |      12    66.98417    8.736285    30.26338    47.75573     86.2126
---------+--------------------------------------------------------------------
combined |      30    56.71033    5.966603    32.68043    44.50726    68.91341
---------+--------------------------------------------------------------------
    diff |           -17.12306    11.96499               -41.63223    7.386123
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.4311
Ho: diff = 0                                     degrees of freedom =       28

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0817         Pr(|T| > |t|) = 0.1635          Pr(T > t) = 0.9183
```

```
-------------------------------------------------------------------------------------------------------------------
-> NewvsOngoingSMDummy = Ongoing Scrum Masters

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |      98    54.92296     3.10403    30.72833    48.76232     61.0836
Treatmen |      83    61.85157    3.259806    29.69825    55.36677    68.33636
---------+--------------------------------------------------------------------
combined |     181    58.10017    2.257723    30.37456    53.64516    62.55518
---------+--------------------------------------------------------------------
    diff |           -6.928607    4.514063               -15.83623    1.979019
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.5349
Ho: diff = 0                                     degrees of freedom =      179

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0633         Pr(|T| > |t|) = 0.1266          Pr(T > t) = 0.9367

.
. sort CompliantTeamSize

. by CompliantTeamSize: ttest PercentageWithTest if Attrition!=1, by(Treatment)

-------------------------------------------------------------------------------------------------------------------
-> CompliantTeamSize = Teams with 12 or more members

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |       1         100           .           .           .           .
Treatmen |       5      46.536    10.15113    22.69862    18.35195    74.72005
---------+--------------------------------------------------------------------
combined |       6    55.44667           .           .           .           .
---------+--------------------------------------------------------------------
    diff |              53.464           .                       .           .
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =        .
Ho: diff = 0                                     degrees of freedom =        4

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =      .         Pr(|T| > |t|) =      .          Pr(T > t) =      .

-------------------------------------------------------------------------------------------------------------------
-> CompliantTeamSize = Teams with 11 or fewer members

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |     115     53.7387    2.878747    30.87113    48.03592    59.44147
Treatmen |      90    63.38678    3.145806    29.84373    57.13613    69.63743
---------+--------------------------------------------------------------------
combined |     205    57.97444    2.146101    30.72749    53.74306    62.20582
---------+--------------------------------------------------------------------
    diff |           -9.648082    4.281903               -18.09079   -1.205373
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -2.2532
Ho: diff = 0                                     degrees of freedom =      203

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0127         Pr(|T| > |t|) = 0.0253          Pr(T > t) = 0.9873

.
. sort CompliantProjectScope

. by CompliantProjectScope: ttest PercentageWithTest if Attrition!=1, by(Treatment)

-------------------------------------------------------------------------------------------------------------------
-> CompliantProjectScope = Teams with 8 or more Features

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |      14    38.17643    7.073711     26.4674    22.89461    53.45825
Treatmen |       9    52.92111    9.547244    28.64173    30.90513     74.9371
---------+--------------------------------------------------------------------
combined |      23    43.94609    5.772464    27.68377    31.97473    55.91745
---------+--------------------------------------------------------------------
    diff |           -14.74468    11.67073               -39.01529    9.525922
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.2634
Ho: diff = 0                                     degrees of freedom =       21
```

```
    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1101         Pr(|T| > |t|) = 0.2203          Pr(T > t) = 0.8899

-------------------------------------------------------------------------------------------------------------------
-> CompliantProjectScope = Teams with 7 or fewer Features

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |     102    56.32824    3.078027    31.08655    50.22226    62.43421
Treatmen |      86    63.50233     3.20735    29.74374    57.12525     69.8794
---------+--------------------------------------------------------------------
combined |     188       59.61      2.2324    30.60913    55.20607    64.01393
---------+--------------------------------------------------------------------
    diff |            -7.17409     4.46219               -15.9771    1.628918
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -1.6078
Ho: diff = 0                                     degrees of freedom =      186

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0548         Pr(|T| > |t|) = 0.1096          Pr(T > t) = 0.9452

.
. /*OUTLIERS ANALYSIS */
.
. histogram TeamSize if Attrition!=1, width(1) start(0) fraction normal fcolor(red) gap(8) ytitle(Percentag
> e of total Teams ) ylabel (,angle(horizontal)) xtitle(Number of Team Members) title("Distributon of Scrum
> s according to Team Size", span size(small)) subtitle( ,ÄùSalience in Recogntion Programmes,Äù, span size(sma
> ll)) note(As of the sixth week of the Quarter)  xlabel(0(1)20)
(bin=22, start=0, width=1)

.
. histogram TotalFeatures if Attrition!=1, width(1) start(0) fraction normal fcolor(red) gap(8) ytitle(Perc
> entage of total Teams ) ylabel (,angle(horizontal)) xtitle(Number of Features per Team) title("Distributo
> n of Scrums according to Total Features", span size(small)) subtitle( ,ÄùSalience in Recogntion Programmes,Äù
> , span size(small)) note(As of the 20th of November 2020)  xlabel(0(1)20)
(bin=16, start=0, width=1)

.
. /*PARAMETRIC TEST OF THE NULL HYPOTHESIS EXCLUDING OUTLIERS */
.
. ttest PercentageWithTest if Attrition!=1 & TeamSize<=11 & TotalFeatures<=13, by(Treatment)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Control |     114    53.66184    2.903076    30.99637    47.91032    59.41336
Treatmen |      90    63.38678    3.145806    29.84373    57.13613    69.63743
---------+--------------------------------------------------------------------
combined |     204    57.95225    2.156532    30.80143    53.70018    62.20433
---------+--------------------------------------------------------------------
    diff |           -9.724936     4.29986               -18.2033   -1.246569
------------------------------------------------------------------------------
    diff = mean(Control) - mean(Treatmen)                         t =  -2.2617
Ho: diff = 0                                     degrees of freedom =      202

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0124         Pr(|T| > |t|) = 0.0248          Pr(T > t) = 0.9876

.
. /* GENERAL LINEAR REGRESSION EXCLUDING OUTLIERS*/
.
. regress PercentageWithTest Treatment if Attrition !=1 & TeamSize<=11 & TotalFeatures<=13

      Source |       SS           df       MS      Number of obs   =       204
-------------+----------------------------------   F(1, 202)       =      5.12
       Model |  4756.53468         1  4756.53468   Prob > F        =    0.0248
    Residual |  187835.278       202  929.877616   R-squared       =    0.0247
-------------+----------------------------------   Adj R-squared   =    0.0199
       Total |  192591.813       203  948.728144   Root MSE        =    30.494

------------------------------------------------------------------------------
Percentage~t |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   Treatment |   9.724936    4.29986     2.26   0.025     1.246569     18.2033
       _cons |   53.66184   2.856015    18.79   0.000     48.03042    59.29327
------------------------------------------------------------------------------

.
. /* GENERAL MULTIPLE LINEAR REGRESSION EXCLUDING OUTLIERS*/
.
. regress PercentageWithTest Treatment TotalFeatures if Attrition !=1 & TeamSize<=11 & TotalFeatures<=13

      Source |       SS           df       MS      Number of obs   =       204
-------------+----------------------------------   F(2, 201)       =     11.06
       Model |  19094.9249         2  9547.46243   Prob > F        =    0.0000
```

```
    Residual |  173496.888       201  863.168599   R-squared       =    0.0991
-------------+----------------------------------   Adj R-squared   =    0.0902
       Total |  192591.813       203  948.728144   Root MSE        =     29.38

------------------------------------------------------------------------------
PercentageW~t |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
    Treatment |   8.815875   4.148755     2.12   0.035     .6352095    16.99654
TotalFeatures |  -4.342159   1.065377    -4.08   0.000    -6.442909    -2.24141
        _cons |   75.22028   5.962425    12.62   0.000     63.46336    86.97721
------------------------------------------------------------------------------

.
. /* A POSTERIORI SAMPLE SIZE ESTIMATION */
.
. sampsi 54.13 62.49, sd1(31.04) sd2(29.66) alpha(0.05) power(0.8)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                  and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            m1 =    54.13
            m2 =    62.49
           sd1 =    31.04
           sd2 =    29.66
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =      207
            n2 =      207

.
. /* TESTS TO VERIFY SAMPLES ARE NOT DIFFERENT AFTER ATTRITION*/
.
. swilk Rank

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
        Rank |        232    0.95947      6.884     4.472    0.00000

. ranksum Rank if Attrition!=1, by(Treatment)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

   Treatment |      obs    rank sum    expected
-------------+---------------------------------
Control Grou |      116       12230       12296
Treatment Gr |       95       10136       10070
-------------+---------------------------------
    combined |      211       22366       22366

unadjusted variance   194686.67
adjustment for ties       -0.00
                     ----------
adjusted variance     194686.67

Ho: Rank(Treatm~t==Control Group) = Rank(Treatm~t==Treatment Group)
             z =  -0.150
    Prob > |z| =   0.8811

. return list

scalars:
                r(N_2) =  95
                r(N_1) =  116
              r(Var_a) =  194686.6666666666
                  r(z) =  -.1495807928433497
            r(sum_exp) =  12296
            r(sum_obs) =  12230
             r(group1) =  0

.
. swilk NewDummy

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
    NewDummy |        232    0.97206      4.745     3.609    0.00015

. tab NewDummy Treatment if Attrition!=1, chi2
```

```
                |       Treatment
        NewDummy | Control G  Treatment |     Total
-----------------+----------------------+----------
Ongoing Projects |        95          82 |       177
    New Projects |        21          13 |        34
-----------------+----------------------+----------
           Total |       116          95 |       211

          Pearson chi2(1) =   0.7546   Pr = 0.385

.
. swilk Q3Score

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
     Q3Score |        232    0.74053     44.068     8.775    0.00000

. ranksum Q3Score if Attrition!=1, by(Treatment)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

   Treatment |      obs    rank sum    expected
-------------+---------------------------------
Control Grou |      116     12008.5       12296
Treatment Gr |       95     10357.5       10070
-------------+---------------------------------
    combined |      211       22366       22366

unadjusted variance   194686.67
adjustment for ties    -4161.29
                     ----------
adjusted variance     190525.38

Ho: Q3Score(Treatm~t==Control Group) = Q3Score(Treatm~t==Treatment Group)
             z =  -0.659
    Prob > |z| =   0.5101

. return list

scalars:
                r(N_2) =  95
                r(N_1) =  116
              r(Var_a) =  190525.3793726021
                 r(z) =  -.6586602148710144
            r(sum_exp) =  12296
            r(sum_obs) =  12008.5
             r(group1) =  0

.
. swilk TotalFeatures

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
TotalFeatu~s |        232    0.92624     12.528     5.860    0.00000

. ranksum TotalFeatures if Attrition!=1, by(Treatment)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

   Treatment |      obs    rank sum    expected
-------------+---------------------------------
Control Grou |      116       12462       12296
Treatment Gr |       95        9904       10070
-------------+---------------------------------
    combined |      211       22366       22366

unadjusted variance   194686.67
adjustment for ties    -5212.05
                     ----------
adjusted variance     189474.61

Ho: TotalF~s(Treatm~t==Control Group) = TotalF~s(Treatm~t==Treatment Group)
             z =   0.381
    Prob > |z| =   0.7029

. return list

scalars:
                r(N_2) =  95
                r(N_1) =  116
              r(Var_a) =  189474.6120514557
                 r(z) =  .3813577483996357
            r(sum_exp) =  12296
```

```
           r(sum_obs) =  12462
            r(group1) =  0

.
. swilk TeamSize

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
    TeamSize |        232    0.96393      6.126     4.201    0.00001

. ranksum TeamSize if Attrition!=1, by(Treatment)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

   Treatment |      obs    rank sum    expected
-------------+---------------------------------
Control Grou |      116     12431.5       12296
Treatment Gr |       95      9934.5       10070
-------------+---------------------------------
    combined |      211       22366       22366

unadjusted variance   194686.67
adjustment for ties    -3042.75
                     ----------
adjusted variance     191643.92

Ho: TeamSize(Treatm~t==Control Group) = TeamSize(Treatm~t==Treatment Group)
            z =   0.310
    Prob > |z| =   0.7569

. return list

scalars:
                r(N_2) =  95
                r(N_1) =  116
              r(Var_a) =  191643.9180771835
                  r(z) =  .3095221793449351
            r(sum_exp) =  12296
            r(sum_obs) =  12431.5
             r(group1) =  0

.
. swilk NewvsOngoingSMDummy

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
NewvsOngoi~y |        232    0.97315      4.560     3.517    0.00022

. tab NewvsOngoingSMDummy Treatment if Attrition!=1, chi2

                     |       Treatment
  NewvsOngoingSMDummy | Control G  Treatment |     Total
----------------------+----------------------+----------
    New Scrum Masters |        18         12 |        30
Ongoing Scrum Masters |        98         83 |       181
----------------------+----------------------+----------
                Total |       116         95 |       211

          Pearson chi2(1) =   0.3566   Pr = 0.550

.
. swilk TrainingDummy

                  Shapiro-Wilk W test for normal data

    Variable |        Obs       W           V         z       Prob>z
-------------+--------------------------------------------------------
TrainingDu~y |        232    0.95497      7.649     4.716    0.00000

. tab TrainingDummy Treatment if Attrition!=1, chi2

                     |       Treatment
        TrainingDummy | Control G  Treatment |     Total
----------------------+----------------------+----------
Untrained Scrum Maste |        12         13 |        25
Trained Scrum Masters |       104         82 |       186
----------------------+----------------------+----------
                Total |       116         95 |       211

          Pearson chi2(1) =   0.5576   Pr = 0.455

.
.
. log close
```

```
      name:  <unnamed>
       log:  /Users/jcvalverde/Dropbox/LSE Behavioural Science/15-PB451E Dissertation/3-Final Intervention/
> 6) Do and txt Files/18080 Dissertation FINAL .txt
  log type:  text
 closed on:  13 Dec 2020, 10:07:46
-------------------------------------------------------------------------------------------------------------
```

9. Bibliography

Alary, D., Gollier, C., & Treich, N. (2013). The Effect of Ambiguity Aversion on Insurance and Self-protection. *The Economic Journal (London)*, *123*(573), 1188–1202.

Appel, V. (1971). Advertising wear out. *Journal of Advertising Research*, *11*(1), 11–13.

Ards, S., Chung, C., & Myers Jr, S. L. (1998). The effects of sample selection bias on racial differences in child abuse reporting. *Child Abuse & Neglect*, *22*(2), 103–115.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, Albert. (1969). *Principles of behavior modification.*

Banerjee, A. V., Banerjee, A., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Bhanot, S., Kraft-Todd, G. T., Rand, D. G., & Yoeli, E. (2018). Putting Social Rewards and Identity Salience to the Test: Evidence from a Field Experiment on Teachers in Philadelphia. *Journal of Behavioral Public Administration*.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803–843.

Bradler, C., Dur, R., Neckermann, S., & Non, A. (2016). Employee recognition and performance: A field experiment. *Management Science*, *62*(11), 3085–3099.

Broadbent, D. E. (1958). Effect of noise on an “intellectual” task. *The Journal of the Acoustical Society of America*, *30*(9), 824–827.

Brown, J., Hossain, T., & Morgan, J. (2010). Shrouded attributes and information suppression: Evidence from the field. *The Quarterly Journal of Economics*, *125*(2), 859–876.

Busse, M., Silva-Risso, J., & Zettelmeyer, F. (2006). $1,000 cash back: The pass-through of auto manufacturer promotions. *American Economic Review*, *96*(4), 1253–1270.

Cacioppo, J. T., & Petty, R. E. (1980). Persuasiveness of communications is affected by exposure frequency and message quality: A theoretical and empirical analysis of persisting attitude change. *Current Issues and Research in Advertising*, *3*(1), 97–122.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research*, *9*(3), 240–244.

Campbell, M. K., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomised trials. *BMJ*, *328*(7441), 702–708. https://doi.org/10.1136/bmj.328.7441.702

Cherry, K. (2020). *How Does the Recency Effect Influence Memory?* Verywell Mind. https://www.verywellmind.com/the-recency-effect-4685058

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145–77.

Cohen, R. T., & Kahana, M. J. (2019). Retrieved-context theory of memory in emotional disorders. *BioRxiv*, 817486.

Conboy, K., Coyle, S., Wang, X., & Pikkarainen, M. (2011). People over Process: Key Challenges in Agile Development. *IEEE Software*, *28*. https://doi.org/10.1109/MS.2010.132

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223.

Cummings, T. G., & Worley, C. G. (2014). *Organization development and change*. Cengage learning.

Darby, S. (2006). *The effectiveness of feedback on energy consumption. A Review for DEFRA of the Literature on Metering Billing and direct Displays,. 486*, 26.

Daunt, V., & Menzies, V. (2020). Recognition programmes: Are they important? [Public]. *Deloitte Perspectives*. https://www2.deloitte.com/ie/en/pages/deloitte-private/articles/recognition-programmes.html

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance*, *64*(2), 709–749.

Dichter, S. F., Gagnon, C., & Alexander, A. (2019). Leading organizational transformations. *McKinsey Quarterly*, *Online Editions*(Business Functions).

Dobbin, F., & Kalev, A. (2016). Why diversity programmes fail. *Harvard Business Review*, *July-August*, 52–61.

Dolan, P., Foy, C., & Smith, S. (2016). The SALIENT checklist: Gathering up the ways in which built environments affect what we do and how we feel. *Buildings*, *6*(1), 9.

Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., & Vlaev, I. (2012). Influencing behaviour: The mindspace way. *Journal of Economic Psychology*, *33*(1), 264–277.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, *56*(293), 52–64.

Dweck, C. S. (2006). *Mindset: The new psychology of success* (1st ed.). Random House.

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). *Retention and obliviscence as a function of the time.*

Edwards, S. J. L., Braunholtz, D. A., Lilford, R. J., & Stevens, A. J. (1999). Ethical issues in the design and conduct of cluster randomised controlled trials. *BMJ*, *318*(7195), 1407–1409. https://doi.org/10.1136/bmj.318.7195.1407

Edwards, W. H. (2010). *Motor learning and control: From theory to practice.* Cengage Learning.

Ellison, G., & Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, *77*(2), 427–452.

Feinberg, C. (2015). The Science of Scarcity. A behavioral economists fresh perspectives on poverty. *Harvard Magazine*, *117*, 38–43.

Galizzi, M. M. (2020). *Research Methods for Behavoural Science—Lecture 6 [Class Handouts]*. Department of Behavioural and Psychological Science, LSE, London England. https://moodle.lse.ac.uk/course/view.php?id=6668

Gallagher, K. S., & Muehlegger, E. (2011). Giving green to get green? Incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, *61*(1), 1–15.

Gallo, A. (2017, June 28). A Refresher on A/B Testing. *Harvard Business Review*. https://hbr.org/2017/06/a-refresher-on-ab-testing

Gimotty, P. A., Burack, R. C., & George, J. A. (2002). Delivery of preventive health services for breast cancer control: A longitudinal view of a randomized controlled trial. *Health Services Research*, *37*(1), 65–86.

Glennerster, R., & Takavarasha, K. (2013). *Running randomized evaluations: A practical guide*. Princeton University Press.

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341), 341ps12–341ps12.

Haak, T. (2013). Do you choose A or B? HR needs more A/B testing. *HR Trend Institute*. https://hrtrendinstitute.com/2013/11/17/do-you-choose-a-or-b-hr-needs-more-ab-testing/

Harper, H. (2019). *Applying Behavioural Insights*. Behavioural Science and Policy Course Guest Lecture, London, England.

Harrison, G. W., & List, Jo. (2004). Field Experments. *Journal of Economic Literature*, *52*(December), 1009–1055.

Hass, D. (2015). *This Is How the Brain Filters Out Unimportant Details* [Public]. Psychology Today. https://www.psychologytoday.com/blog/brain-babble/201502/is-how-the-brain-filters-out-unimportant-details

Hayes, M. J., Svoboda, M. D., Wilhite, D. A., & Vanyarkho, O. V. (1999). Monitoring the 1996 drought using the standardized precipitation index. *B. Am. Meteorol. Soc.*, *80*, 429–438.

Haynes, L., Goldacre, B., Torgerson, D., & others. (2012). Test, learn, adapt: Developing public policy with randomised controlled trials. *Cabinet Office-Behavioural Insights Team*.

Heath, T. B., & Chatterjee, S. (1995). Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence. *Journal of Consumer Research*, *22*(3), 268–284.

Hesselberg, J. (2018). *Unlocking Agility: An Insider's Guide to Agile Enterprise Transformation*. Addison-Wesley Professional.

Higgins, E. T. (1996). Activation: Accessibility, and salience. *Social Psychology: Handbook of Basic Principles*, 133–168.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion.*

Jost, A. (1897). *Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen*. L. Voss.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401–405.

Kahneman, D., & Tversky, A. (2000). Choices, values, and frames. In *<h>Choices</h>, <h>values</h>, <h>and frames</h>*. Cambridge University Press.

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., Fernández, M. E., Markham, C., & Bartholomew, L. K. (2015). A taxonomy of behavior change methods; an intervention mapping approach. *Health Psychology Review*. https://doi.org/10.1080/17437199.2015.1077155

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall International.

Komaki, J., Coombs, T., & Schepman, S. (1996). Motivational implications of reinforcement theory. *Motivation and Leadership at Work*, *34*, 52.

Komaki, J. L., Coombs, T., & Schepman, S. (1996). *Motivational implications of reinforcement theory. Published in RM Steers, LW Porter, & GA Bigley (eds.) Motivation and leadership at work (pp. 34-52)*. New York: McGraw-Hill.

Kompella, L. (2014). Agile Methods, Organizational Culture and Agility: Some Insights. *Proceedings of the 7th International Workshop on Cooperative and Human Aspects of Software Engineering*.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-Five Years After Broadbent (1958): Still No Identification Without Attention. *Psychological Review*, *111*(4), 880–913.

Landry, C. E., Lange, A., List, J. A., Price, M. K., & Rupp, N. G. (2006). Toward an understanding of the economics of charity: Evidence from a field experiment. *The Quarterly Journal of Economics*, *121*(2), 747–782.

Latham, G. P. (2012). *Work motivation: History, theory, research, and practice*. Sage.

List, J., & Gneezy, U. (2014). *The why axis: Hidden motives and the undiscovered economics of everyday life*. Random House.

Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 12.

Lorenc, T., Petticrew, M., Welch, V., & Tugwell, P. (2013). What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health*, *67*(2), 190–193.

Luthans, F., & Stajkovic, A. D. (2000). Provide recognition for performance improvement. *Handbook of Principles of Organizational Behavior*, 239.

Luthans, K. (2000). Recognition: A powerful, but often overlooked, leadership tool to improve employee performance. *Journal of Leadership Studies*, *7*(1), 31–39.

Mack, C. C., Cinel, C., Davies, N., Harding, M., & Ward, G. (2017). Serial position, output order, and list length effects for words presented on smartphones over very long intervals. *Journal of Memory and Language*, *97*, 61–80.

McCullough, J. L., & Ostrom, T. M. (1974). Repetition of highly similar messages and attitude change. *Journal of Applied Psychology*, *59*(3), 395.

Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *The Journal of Abnormal and Social Psychology*, *59*(1), 1.

Miller, N. E., & Stevenson, S. S. (1936). Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery. *Journal of Comparative Psychology*, *21*(2), 205.

Mitchell, T. R., & Mickel, A. E. (1999). The meaning of money: An individual-difference perspective. *Academy of Management Review*, *24*(3), 568–578.

Moran, J. W., & Brightman, B. K. (2000). Leading organizational change. *Journal of Workplace Learning*.

Morse, G. (2016). Designing a bias-free organization. *Harvard Business Review*, *94*(7), 15.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482.

Murray, D. M., Taljaard, M., Turner, E. L., & George, S. M. (2020). Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials. *Annual Review of Public Health*, *41*(1), 1–19. https://doi.org/10.1146/annurev-publhealth-040119-094027

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*(4), 480–494.

Nickerson, D. W. (2007). Quality Is Job One: Professional and Volunteer Voter Mobilization Calls. *American Journal of Political Science*, *51*(2), 269–282.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606.

Petrosino, A., Turpin-Petrosino, C., & Finckenauer, J. O. (2000). Well-meaning programs can have harmful effects! Lessons from experiments of programs such as Scared Straight. *Crime & Delinquency*, *46*(3), 354–379.

Pieters, R., & Bijmolt, T. (1997). Consumer Memory for Television Advertising: A Field Study of Duration, Serial Position, and Competition Effects. *Journal of Consumer Research*, *23*, 362–372. https://doi.org/10.1086/209489

Progress Corp. (2013). *The Importance of Prioritizing and Sizing Your Backlog for A*. Telerik Blogs. https://www.telerik.com/blogs/the-importance-of-prioritizing-and-sizing-your-backlog-for-agile-planning

Quirke, M. B. (2012). *Making the connections: Using internal communication to turn strategy into action*. Gower Publishing, Ltd.,.

Salameh, H. (2014). What, when, why, and how? A comparison between agile project management and traditional project management methods. *International Journal of Business and Management Review*, *2*(5), 52–74.

Scaled Agile Inc. (2019). Team and Technical Agility [Public]. *Public*. https://www.scaledagileframework.com/team-and-technical-agility/

Schippers, M. C., Scheepers, A. W., & Peterson, J. B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications*, *1*(1), 1–12.

Sexton, S. (2015). Automatic Bill Payment and Salience Effects: Evidence from Electricity Consumption. *The Review of Economics and Statistics*, *97*(2), 229–241.

Shrestha, P. (2017). Ebbinghaus Forgetting Curve. *Psychestudy*. https://www.psychestudy.com/cognitive/memory/ebbinghaus-forgetting-curve

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Skinner, B. F. (2019). *The behavior of organisms: An experimental analysis*. BF Skinner Foundation.

Spacey, J. (2017). *Four Examples of Recency*. Simplicable. https://simplicable.com/new/recency

Stajkovic, A. D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance, 1975–95. *Academy of Management Journal*, *40*(5), 1122–1149.

Stajkovic, A. D., & Luthans, F. (2001). Differential effects of incentive motivators on work performance. *Academy of Management Journal*, *44*(3), 580–590.

Stajkovic, A. D., & Luthans, F. (2003). Behavioral management and task performance in organizations: Conceptual background, meta-analysis, and test of alternative models. *Personnel Psychology*, *56*(1), 155–194.

Stewart, N. (2009). The Cost of Anchoring on Credit-Card Minimum Repayments. *Psychological Science*, *20*(1), 39–41.

Stewart, S., Rinke, E. M., McGarrigle, R., Lynott, D., Lunny, C., Lautarescu, A., Galizzi, M. M., Farran, E., & Crook, Z. (2020). *Pre-registration and Registered Reports: A primer from UKRN*.

Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, *56*, 65.

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive" vividness" effect. *Psychological Review*, *89*(2), 155.

Terry, W. S. (2005). Serial position effects in recall of television commercials. *The Journal of General Psychology*, *132*(2), 151–164.

Tesser, A. (1986). Some effects of self-evaluation maintenance on cognition and action. In Sorrentino & Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behaviour*. Guilford Press.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In *Advances in experimental social psychology* (Vol. 21, pp. 181–227). Elsevier.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thorndike, E. L. (1913). *The psychology of learning* (Vol. 2). Teachers College, Columbia University.

Vlaev, I., & Dolan, P. (2009). *From changing cognitions to changing the context: A dual-route model of behaviour change*.

Weik, K. E. (1995). *Sensemaking in organisations* (Vol. 3). Sage.

Wells, K. (2017). *3 KPIs For Measuring Internal Communication Effectiveness*. Field Service Digital. https://fsd.servicemax.com/2017/10/04/measuring-internal-communication-effectiveness/

Zenger, J., & Folkman, J. (2014, January 15). Your Employees Want the Negative Feedback You Hate to Give [Public]. *Harvard Business Review Online*. https://hbr.org/2014/01/your-employees-want-the-negative-feedback-you-hate-to-give