The Leadership Quarterly xxx (2022) 101630

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Full length article

The pandemic that shocked managers across the world: The impact of the COVID-19 crisis on leadership behavior

Harry Garretsen[a], Janka I. Stoker[a,*], Dimitrios Soudis[a], Hein Wendt[b,c]

[a] University of Groningen, Faculty of Economics and Business, the Netherlands
[b] Korn Ferry, Amsterdam, the Netherlands
[c] KU Leuven, Belgium

ARTICLE INFO

Keywords:
COVID-19
Differences-in-differences
Directive leadership
Participative leadership
Threat-rigidity hypothesis
Crisis
Exogenous shocks
Working From Home

ABSTRACT

In March 2020, the COVID-19 virus turned into a pandemic that hit organizations globally. This pandemic qualifies as an exogenous shock. Based on the threat-rigidity hypothesis, we hypothesize that this shock led to an increase in directive leadership behavior. We also argue that this relationship depends on the magnitude of the crisis and on well-learned responses of managers. In our empirical analysis we employ a differences-in-differences design with treatment intensity and focus on the period of the first lockdown, March until June 2020. Using a dataset covering monthly data for almost 27,000 managers across 48 countries and 32 sectors for January 2019 to December 2020, we find support for the threat-rigidity hypothesis. During the first lockdown, directive leadership increased significantly. We also find that this relationship is moderated by COVID-19 deaths per country, the sectoral working from home potential, and the organizational level of management. Our findings provide new evidence how large exogenous shocks like COVID-19 can impact leadership behavior.

Introduction

The COVID-19 pandemic is an exogenous shock that affects individuals and organizations all over the world. Notwithstanding the terrible and devastating effects of this major health crisis, the pandemic offers quite a unique research opportunity for leadership scholars like ourselves, because it enables us to study the possible effects of such a systemic global shock on leadership behavior. Exogenous shocks can be defined as sudden changes that can dramatically affect individuals, organizations, and society at large (Meyer, 1982; Ramey, 2016). Studying the impact of such shocks on leadership behavior helps to advance our field in multiple ways. First and foremost, exogenous shocks allow for the inference of causal relationships. As illustrated by the excellent review of Sieweke and Santoni (2020), the leadership field is taking important and necessary steps to advance causal research (Antonakis, 2017; Antonakis, Bendahan, Jacquart, & Lalive, 2010), e.g., by casting exogenous shocks to study the impact on leadership behavior (Stoker, Garretsen, & Soudis, 2019).

Second, exogenous shocks create the possibility to rigorously investigate the relevance and impact of context on leadership. Although the importance of the context is rather undisputed in the leadership field (Oc, 2018; Yukl, 2013), the way how to conceptualize and measure relevant contextual variables is not (Garretsen, Stoker, & Weber, 2020). Contextual variables are typically included as a moderator in the relationship between leadership and outcome variables, whereas the context can also be an antecedent of leadership behavior (Stoker et al., 2019; Tuncdogan, Acar, & Stam, 2017). Moreover, although several leadership and management scholars recognize that the context should also be studied at meso- and macro levels (Aguinis, Boyd, Pierce, & Short, 2011; Hiller, Piccolo, & Zaccaro, 2020; Johns, 2018; Oc, 2018), leadership research has characteristically defined and measured context at the micro-level of the individual leader with his/her follower(s), with a few exceptions (e.g., Jacquart & Antonakis, 2015; Stoker et al., 2019).

By studying the impact of the COVID-19 pandemic on leadership behavior, our article contributes to the advancement of the field where it concerns both causal research and the role of context. The COVID-19 shock is obviously a very relevant system-wide exogenous shock, one which impacts on individuals, organizations, sectors and countries globally (Jacquart, Santoni, Schudy, Sieweke, & Withers (2020)). Our study builds on earlier work in which the

* Corresponding author at: University of Groningen, Faculty of Economics and Business, PO Box 800, 9700 AV Groningen, The Netherlands.
E-mail address: j.i.stoker@rug.nl (J.I. Stoker).

https://doi.org/10.1016/j.leaqua.2022.101630
Received 15 March 2021; Revised 27 May 2022; Accepted 2 June 2022
Available online xxxx

impact of the 2008 financial crisis on leadership behavior was investigated (Stoker et al., 2019). Based on the threat-rigidity hypothesis, which argues that individuals and organizations react to an external threat with actions that reflect rigidity (Staw, Sandelands, & Dutton, 1981), Stoker et al. (2019) show for the 2008 financial crisis that it went along with a significant increase in directive leadership, and that this effect was context-dependent. Following calls for more studies that engage in replications in leadership research (Antonakis, 2017), the present study attempts to show whether or not the main findings of how large systemic shocks impact on leadership behavior are replicated, by comparing the findings for the 2008 financial crisis with the COVID-19 pandemic. At the same time, the current study extends earlier work by using a broader set of contextual variables and also by employing a more fine-grained data set.

Following threat-rigidity literature and empirical findings (Staw et al., 1981, Stoker et al., 2019), we expect that the COVID-19 shock is associated with an increase in directive leadership. In addition, and by making use of the richness of our data set, we investigate the moderating effect of a range of contextual variables at multiple levels on the relationship between the COVID-19 shock and leadership behavior. Based on the threat-rigidity hypothesis (Staw et al., 1981), these contextual variables relate either to the magnitude of the crisis, or to the relevance of so-called 'well-learned responses'. At the country- or macro- level, we use two contextual variables. We cast the number of COVID-19 deaths as an indicator of the magnitude of the crisis. As indicator of well-learned responses at the country-level, we include power distance (see also Stoker et al., 2019). By way of robustness check, we will also check whether and how participative leadership changed because of the COVID-19 shock, the idea being that the threat-rigidity hypothesis would not predict an increase in participative leadership as well.

At the sectoral- or meso-level, we use data on the working from home potential (WFHP). The WFHP crucially predates the COVID-19 shock and thus signals, irrespective of the COVID-19 crisis, how many jobs or tasks within a certain sector of the economy could feasibly be done by working from home (Dingel & Neiman, 2020). It captures per sector how many jobs or tasks potentially could be done via WFH, given the technological or organizational requirements of the job or task at hand. WFHP is therefore not a measure of the actual working from home that followed in the wake of the COVID-19 shock. For sectors in which the WFHP is low (high), we expect that well-learned responses by managers will be more (less) relevant as a reaction to the COVID-19 shock.

Finally, at the organizational- or micro-level, we include the level of management as a moderating factor. Research shows that directive leadership tends to more dominant at lower management levels (Lowe, Kroeck, & Sivasubramaniam, 1996; Oshagbemi & Gill, 2004), which indicates that this behavior is a well-learned response at these lower levels.

We use a large, longitudinal database containing monthly data on subordinates' assessments of their manager's behaviors (see Euwema, Wendt, & Van Emmerik, 2007; Stoker et al., 2019; Van Emmerik, Wendt, & Euwema, 2010). Because the number and the identity of the managers that are being assessed vary across time, our data allow for a repeated cross-sectional analysis. Since the COVID-19 crisis affects potentially all organizations and managers across the globe, a standard differences-in-differences (DID) experiment with a treatment and control group is obviously not feasible. So, how to deal with global shocks like the COVID-19 crisis?

In the framing of a DID-design – and hence instead of a standard treatment–control classification – the alternative strategy is to exploit the heterogeneity in the data, so as to come up with classification of groups (in our case, groups of managers and their organizations) that vary in the 'treatment intensity' of the shock. In economics (Angrist & Pischke, 2009, chapter 5), but also in management and leadership research, this approach with a treatment intensity variable is quite common as an alternative DID-research design (St. Clair & Cook, 2015). In our study, we will use the idea that managers can be classified in relative treatment terms as to how the COVID-19 shock might have affected their behavior to set up our hypotheses. This relative treatment, or treatment intensity, of the COVID-19 shock will be captured by the moderating variables mentioned above, namely the WFHP per sector, the relative number of COVID-19 deaths per country, and the level of management per organization.

Next to the choice for a DID-research design with treatment intensity to study the exogenous shock that is the COVID-19 crisis, the demarcation of the period for which such a shock can be qualified as truly exogenous is crucial. We will argue that the starting date of the exogenous shock can be determined as March 1st 2020. Once the COVID-19 crisis had started in March 2020, the exogeneity of the shock wears off after some time, because policy and other (leadership) interventions start to affect the crisis, such that invariably at some point endogeneity creeps in. Therefore, in our analysis we have to be careful about the period for which the shock can be deemed to be exogenous.

In our view, and as we will explain in section 3, the exogeneity of the COVID-19 crisis is safeguarded most when we restrict the crisis period to the period of the first lockdown, that is from March until June 2020. Next to this period, we will also conduct our analysis for our full crisis sample period, March to December 2020. Using a selection of the aforementioned database covering monthly data for the period January 2019 to December 2020 for in total almost 27,000 managers in 633 organizations across 48 countries and 32 sectors, we compare leadership behavior before and after the onset of the COVID-19 pandemic, where we will thus take March 1st 2020 as the demarcation date. We test a multilevel model that incorporates our contextual or moderating variables at three different levels: the organizational, sectoral, and country level.

Our study contributes to leadership and management research in three ways. First, by empirically investigating the effects of an exogenous shock on leadership behaviors, our article is a response to the plea in the field of leadership for more rigorous research (Antonakis et al., 2014; Sieweke & Santoni, 2020). Our findings confirm that leadership behavior is impacted by exogenous shocks. More specifically, and here lies our second contribution, we contribute to the growing literature on crises and specifically on the effect of the COVID-19 crisis on leadership (Kniffin et al., 2021; Rudolph et al., 2021). Finally, by testing a multi-level model in a DID research design setting, we are able to show how the effect of the threat of the COVID-19 pandemic is moderated by contextual variables and how these moderating effects help to establish whether the impact of the COVID-19 on leadership behavior is indeed a causal one. In doing so, we not only contribute to threat-rigidity hypothesis, but also provide support for the claim that it is necessary to include the broader meso- and macro-context in leadership and management research (Aguinis et al., 2011).

As we already stated above, our article builds on previous comparable work where the impact of the 2008 financial crisis was investigated (Stoker et al., 2019). We extend this earlier work in three crucial ways. First and foremost, in the current article we study an exogenous shock which has different and unique characteristics when compared to the financial crisis of 2008. Unlike the financial crisis of 2008, the COVID-19 crisis is first and foremost a health crisis. Also, and quite uniquely, the COVID-19 pandemic led to a situation where many managers and their followers had to work from home (Kniffin et al., 2021, Stoker, Garretsen, & Lammers, 2022). Second, comparing the data on leadership behavior from the 2008 financial crisis article (Stoker et al., 2019) with the current dataset, we are in the fortunate position that we have data on a monthly basis, allowing us to really pinpoint the start of the crisis, but also to measure within-year behavioral changes in the responses of managers after the start of the COVID-19 pandemic. Related, the current dataset also has more information when it comes to the level of management. The study on the financial crisis (Stoker et al., 2019) was not able to investigate possible differences of the effect of the crisis on lower or higher management, while research shows that leadership behavior differs across levels in the hierarchy (Chun, Yammarino, Dionne, Sosik, & Moon, 2009). A final important difference concerns the urgency to learn more about this still unfolding shock for leadership behavior. Instead of looking back on a previous shock episode (like in Stoker et al., 2019), both academics and practitioners need more knowledge of the COVID-19 pandemic in the here and now (see also Kniffin et al., 2021).

Notwithstanding these notable differences with Stoker et al. (2019), we also firmly believe that we can learn much from replications (Antonakis, 2017). In line with the current replication debate (see e.g. Serra-Garcia & Gneezy, 2021), it is crucial to seriously investigate if the main conclusion from Stoker et al. (2019) that a crisis went along with an increase in directive leadership behavior is a one-off, or whether it can be replicated and hence be given more validity by looking at other systemic crises. We see our current study very much as a contribution to a wider research agenda for leadership scholars (Garretsen et al., 2020).

Theoretical background

The COVID-19 crisis and the threat-rigidity hypothesis

In order to determine the possible effects of the COVID-19 crisis on leadership behavior, we follow assumptions from the threat-rigidity hypothesis, in which Staw et al. (1981, p. 502) define a threat as “an environmental event that has impending negative consequences for the entity”. The basic assumption of the threat-rigidity hypothesis is that organizations, and individuals like managers within those organizations, react with rigidity towards a threat. That is, as a response to such a threat they will exercise more control, will restrict information processing and will centralize decision making (see also Gladstein & Reilly, 1985). For the 2008 financial crisis, Stoker et al. (2019) argued that such general managerial responses could be translated into a specific leadership style, that is directive leadership (Kamphuis, Gaillard, & Vogelaar, 2011). Directive leadership can be defined as giving clear and detailed directions to employees, exercising control, and expecting compliance with instructions (see e.g. House, 1971; Kamphuis et al., 2011; Lorinkova, Pearsall, & Sims, 2013; Somech, 2005, 2006). The 2008 financial crisis was associated with a significant increase in directive leadership (Stoker et al., 2019).

The COVID-19 crisis clearly meets the definition of a threat of Staw et al. (1981), both from a health and an economic perspective. This crisis was unprecedented, because it was the largest global health shock since the Spanish Flu of 1918 (Franchini et al., 2020). Although we have had other global health crises over the last decades, like the near pandemic of Severe Acute Respiratory Syndrome (SARS) in 2003, or the widespread pandemic-like extension of Ebola over five African countries (which started around 1976), the current COVID-19 pandemic causes incredible suffering, death, and disruption of normal life, thereby qualifying as an unparalleled existential threat to the world population (Morens, Daszak, Markel, & Taubenberger, 2020).

It was also an unexpected crisis, as is shown by policy measures data (Hale, Petherick, Phillips, & Webster, 2020). Although the first COVID-19 infected patient in Wuhan was diagnosed on December 8th, 2019 (Shangguan, Wang, & Sun, 2020), it took governments throughout the world several months before they responded, particularly ramping up in March 2020 when the first COVID-19 infections and deaths occurred within their own countries (Morens et al., 2020). Finally, the COVID-19 pandemic can not only be regarded as an unprecedented global health crisis, but it also (initially) triggered (expectations about) a worldwide economic crisis (Moorty et al., 2020; IMF, 2020).

As a consequence, and given also that we focus in our analysis on the initial months in 2020 at the start of the pandemic, we argue that the COVID-19 crisis convincingly can be casted as an external threat, as is described in the threat-rigidity hypothesis. Building on findings of leadership responses to other comparable external threats, notably Stoker et al. (2019), we therefore hypothesize that the COVID-19 crisis leads to a situation in which managers are inclined to increase their levels of control. More specifically, our hypothesis 1 is:

Hypothesis 1. The COVID-19 crisis led to an increase in directive leadership.

Hypothesis 1 builds on earlier findings of a comparable global crisis, the 2008 financial crisis, in which a similar effect on directive leadership was observed (Stoker et al., 2019). Just like in Stoker et al. (2019) and by way of robustness test, we also investigate the possible effect of the pandemic on participative leadership based on the expectation that no significant change should occur, according to the threat-rigidity hypothesis. In a similar vein, we will also estimate our models with moderating variables (see below) for participative leadership as well. It is important to stress that these two leadership styles, directive and participative leadership, are not opposite ends of a single continuum. Although several studies find significant negative correlations, these studies also show both styles have distinct effects, and can very well be used by leaders at the same time (Somech, 2005; Stoker et al., 2019).

Moderating variables, the COVID-19 shock and directive leadership

Our follow-up question is whether contextual factors determine if and how directive leadership increases after this exogenous shock. Following threat-rigidity literature (Staw et al., 1981), we focus on the magnitude of the crisis and on so-called well-learned responses (see also Chattopadhyay, Glick, & Huber, 2001), and propose that these two dimensions affect the impact of the external shock on directive leadership. As we stated already in our introduction, we will take moderating factors into account at three different levels: macro (country), meso (sector) and micro (organization).

External threats or crises like COVID-19 vary in their magnitude across countries (Chattopadhyay et al., 2001; Shimizu, 2007; Stoker et al., 2019). Data show that the impact of the pandemic was not evenly spread across countries in terms of COVID-cases and notably COVID-deaths. Some countries managed to keep the number of deaths low, like New-Zealand (27 deaths as of February 2021) or Iceland (29 deaths in February 2021), whereas other countries like the US (>500,000) or Brazil (around 250,000) suffered far more, not only in absolute terms but also in relative terms (deaths as % of total population).[2]

We propose the magnitude of the COVID-19 crisis to be stronger for countries with a relatively higher number of COVID-19 deaths, which will consequently lead to an increase in directive leadership. The basic assumption (Staw et al., 1981; Chattopadhyay et al., 2001) is that more severe threats call for stronger and more controlling organizational and hence managerial actions. We therefore hypothesize:

[2] See: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?%22%20%5Cl%20%22countries.

Hypothesis 2. The magnitude of the COVID-19 crisis, as measured by the relative number of COVID-19 deaths, moderates the relationship between the COVID-19 crisis and directive leadership. The crisis will have a stronger positive effect on directive leadership in countries with relatively more COVID-19 deaths.

According to the threat-rigidity hypothesis (Staw et al., 1981), a threat leads to the use of well-learned responses or dominant routines. Such responses are self-reinforcing, such that in the case of managers, they respond to a threat by relying on and using well-learned behaviors, instead of trying to find new or different forms of behaviors (c.f. Gilbert, 2005). Consequently, the leadership behaviors or styles that are learned or representative to begin with, will be reinforced as an effect of a crisis (Dickson, Den Hartog, & Mitchelson, 2003). We cast the relevance of these well-learned responses as moderating variables on three levels namely macro (country), meso (sector) and micro (level of management). We will discuss each of these variables in turn.

At the macro- or country-level, one of the key determinants of the context in which managers act and 'learn' their behaviors are cultural values (see e.g., the GLOBE studies by House, Hanges, Javidan, Dorfman, & Gupta, 2004)), like power distance (Dickson et al., 2003; Kirkman, Chen, Farh, Chen, & Lowe, 2009; Stoker et al., 2019). Power distance refers to the way in which in general social relationships in a society are perceived to be hierarchical and unequal (Hofstede, 1980; House, Javidan, Hanges, & Dorfman, 2002; House et al., 2004). It has been shown that directive leadership in countries with high power distance is on average higher than in countries with low power distance (Jackson, Meyer, & Wang, 2013; Stoker et al., 2019). Moreover, it has been found that power distance significantly moderated the positive relationship between a crisis and directive leadership. Specifically, the increase of directive leadership after the 2008 financial crisis was higher in countries with high levels of power distance (Stoker et al., 2019). Following these findings, we therefore hypothesize:

Hypothesis 3a. Power-distance moderates the relationship between the COVID-19 crisis and directive leadership. The crisis will have a stronger positive effect on directive leadership under conditions of high power distance than under conditions of low power distance.

With respect to the moderating variable on the meso- or sector-level, we investigate the level of WFHP across sectors, and hence the 'between-sector' variation in WFHP across organizations and managers. Irrespective of the COVID-19 shock, organizations in e.g. the manufacturing or transportation sector have less possibility to send their employees home to do their work, than organizations in financial or other services; this generic possibility to work from home in each sector has been described by Dingel and Neiman (2020) as the potential to work from home (WFHP; see also the section 'data set and model specification' for the WFHP measure used). To be clear, the WFHP variable employed will not measure actual WFH during the COVID-19 crisis, but merely the sectoral differences in the potential to do so. This measure of WFHP variation thus pre-dates the COVID-19 shock and the implications this shock clearly had for actual WFH from March 2020 onwards.

WFHP is a relevant moderating variable, because the potential to be able to work from home affects the interaction between managers and their employees. A crucial difference between WFH and 'normal' working conditions is namely that managers who work from home have to lead their employees from a distance. Therefore, when managers and employees work from home, there is less room for using well-learned leadership responses when confronted with the COVID-19 crisis (Bonet & Salvador, 2017). Moreover, early research on changes in leadership behavior during the COVID-19 crisis confirms that WFH specifically affects the possibility to execute directive or controlling leadership behaviors. For a Danish sample of 1053 employees and 290 managers during the lockdown in the Spring of 2020, Kirchner, Ipsen, and Hansen (2021) found that direct contact and face-to-face communication about work-related tasks became more difficult for managers.

We therefore propose that for managers, the well learned response of directive leadership as a response to a shock depends on the possibility to execute these behaviors. We assume that in an organizational context where the WFHP is low and thus managers and their employees probably will still be doing their jobs in their 'normal' work place, it is more likely to revert to the well-learned response of directive leadership; whereas in contexts where the WFHP is high, such a well learned response of directive leadership is less likely and less possible. We thus hypothesize that:

Hypothesis 3b. WFHP moderates the relationship between the COVID-19 crisis and directive leadership. The crisis will have a stronger positive effect on directive leadership under conditions of low WFHP than under conditions of high WFHP.

Finally, at the micro-level, we expect the impact of the COVID-19 crisis on directive leadership to depend on the level of management. Leadership behaviors vary across hierarchical levels of managers (Chun et al., 2009; Jago & Vroom, 1977). Whereas managers at lower levels tend to give specific directives, higher level managers "give only broad outlines, opinions and suggestions" (Oshagbemi, 2008, p. 1906). Lower level managers have more interactions with employees than higher level managers, making it more likely that directive leadership is the more dominant routine amongst these managers. Previous research has indicated that a directive style of leading tends to be more dominant at lower management levels (Lowe et al., 1996; Oshagbemi & Gill, 2004; Van Emmerik et al., 2010). Following the threat-rigidity hypothesis, we expect that this well-learned response of directive leadership will be stronger for lower level managers. We therefore expect a stronger effect of the COVID-19 crisis on lower level managers, and hypothesize that:

Hypothesis 3c. The level of management moderates the relationship between the COVID-19 crisis and directive leadership. The crisis will have a stronger positive effect on directive leadership for low compared to high levels of management.

Please recall that as a robustness check, and for each of the moderating hypotheses, we will also run our models for participative leadership.

Research design

As stated in the call for papers for this special issue of *The Leadership Quarterly,* exogenous shocks not only create the opportunity to analyze how leadership is affected by such shocks, but also provide a novel ground for testing causal claims in leadership theories. When it comes to the research design employed to study the impact of exogenous shocks on leadership behavior or leadership outcomes, various strategies exist to infer causality (see Antonakis, 2017; Antonakis et al., 2019; Garretsen et al., 2020; see also Sieweke & Santoni, 2020, for a recent review). The use of for instance instrumental variables, event studies, a regression discontinuity analysis (RDD) or a Differences-in-Differences analysis are all examples of well-known 'tools' that are increasingly used in leadership research to this end.

Our main aim is to investigate whether the level of directive leadership behavior shown by managers *on average* differs when comparing managers' directive leadership behavior pre- and post-COVID 19. We thus also use various moderating variables. In doing so, we effectively conduct a Differences-in-Differences (DID) analysis with a treatment intensity variable (see for a survey St. Clair & Cook, 2015, pp. 330–331). In a seminal book on research design and techniques like

DID, Angrist and Pischke (2009) describe a DID set-up which mirrors our model set-up, with moderators like e.g. the sectoral WFHP being identical to their DID-model with similar interaction terms (see Chapter 5, pp. 233–237).[3]

In our case, viable or useful 'treatment intensity' variables are variables that are not themselves impacted by the COVID-19 shock. Our moderating variables qualify as such treatment intensity variables, including notably the WFHP variable. Note again, quite crucially, that this variable does not measure actual level of Working From Home in the wake of COVID-19 crisis (this would create endogeneity issues). Instead it measures, completely irrespective of the COVID-19 shock, the (technological) potential per sector for jobs or tasks to be done from home on the sectoral level (Dingel & Neiman, 2020).

Irrespective of the research design as such, any credible statement about causality, in the sense that a shock could *cause* changes in leadership behavior in our sample, has to meet some specific requirements. To illustrate why we believe we can assume that the COVID-19 shock meets those requirements, we offer the following arguments. First and foremost, the arrival of the COVID-19 pandemic was clearly exogenous to individual organizations and their managers. Second, the shock was sudden and unexpected. There is ample evidence[4] that in the first two months of 2020, politicians, financial markets and also health and economic forecasters were *not* expecting the 'virus from Wuhan' to turn into a pandemic, let alone affect their country in a serious way in the sense that soon 'stay and work at home' policy measures had to be taken on a truly unprecedented scale (IMF, 2020). This timeframe matters for our research design, because the turning point in terms of both the COVID-19 pandemic and the ensuing policy measures, like ordering a large part of the labor force to stay and work at home, can be pinpointed to the month of March 2020.

In March 2020, almost all countries saw the COVID-19 infection and causality rates go from a virtual flat-liner to exponential growth (see Fig. 1 for COVID-deaths per million people for five selected countries; organizations and managers from these countries figure prominently in our own data set). At almost the same instance, and hence within the same month of March 2020, policy measures went from almost none or very few, to far-reaching and society-wide lockdowns or other stringent measures. Fig. 2 displays (again for five countries) the government stringency index, a composite measure of the main government policy measures in reaction to COVID-19, with a strong weight for stay and work at home measures. Figs. 1 and 2 were made by using the data collected by Our World in Data[5] and the related and also publicly available data set of researchers from Oxford University that tracks the government responses to the COVID-19 pandemic, including WFH.[6]

This second requirement, that we must know when the exogenous shock hit our sample of organizations and their managers, is thus fulfilled in the sense that the data clearly show that the month of March 2020 is globally the demarcation period to distinguish between the 'before' and 'after' situation. Fortunately, and crucially, our data on organizations and managers are collected on a monthly basis, such that we can sharply distinguish leadership behavior and other variables before and after March 1st, 2020. A third requirement for a useful analysis of the impact of the COVID-19 shock on leadership behavior is that one must be able to include the potential variation in the impact on leadership across various contexts. To tackle this issue, we use a multi-level model, including three contextual levels.

A fourth and final requirement to deal with the exogeneity issue in our empirical analysis is that we do not only need a starting date (March 1st, 2020), but also a date at which the exogeneity might be compromised. Invariably over time, the initial shock becomes intertwined with policy interventions and other behavioral or man-made responses to the shock. This makes the shock endogenous to at least some degree. It is for this reason that we consider the period of the first lockdown, which can be pinpointed from March until June 2020, to be the most appropriate period to test our hypotheses. In many countries, this first lockdown period with very strict policy measures lasted from March 2020 until June 2020. The reason to distinguish the first three months (March, April, and May 2020) from the remainder of 2020, is that this quarter after the start of the crisis, does not only concern the period of the initial, unexpected wave of COVID-19 infections and deaths, but also covers the period when many organizations started to adapt to the consequences of this external threat. From June 2020 onwards, the strength or severity of the COVID-19 crisis waned off. There is clear evidence for this: in many countries, lockdown measures were relaxed or lifted after the number of COVID-19 infections and deaths came down strongly in the course of 2020, see also Fig. 1.[7]

To sum up, we are dealing with a unique, systemic and truly exogenous shock that characterizes the COVID-19 pandemic. In the setting of our DID-design, this shock can be identified as occurring as of March 1st 2020 across countries (with the exception of China, see below) and can considered to be exogenous for the March until June 2020 period. Crucially, the magnitude and well-learned responses are instrumental in determining the impact of the COVID-19 shock on directive leadership and our data set takes into account four relevant contextual variables at the micro, meso and macro level as indicators of these two dimensions.

Data set and model specification

Sample and procedure

Our primary data source stems from a unique data set collected by Korn Ferry (hereafter: KF), a worldwide operating consulting firm. The KF data collection originates from the assessment of both managers themselves and their subordinates, an assessment that took place before management training programs by KF within each of the participating organizations, which guarantees a very high response rate. The KF data set covers many profit and non-profit organizations across a

[3] The example used in Angrist and Pischke (2009) is a study by one the founding fathers of the use of DID-model in modern labour economics, David Card (Card, 1992), see Angrist and Pischke (2009, pp 234-237). Like we will do in our empirical analysis, Card (1992) also uses a *two period* setting (i.c. before and after a shock or regime change like, in his case, a change of the minimum wage across US states) to investigate the impact on teen employment per US state of the treatment (change minimum wage) interacted with the treatment intensity variable (fraction of teenagers per state to be affected by minimum wage). This idea is exactly the same as our models with interaction effects, where in the absence of a true non-treatment or control group, the treatment intensity variable when interacted with the shock enables the testing of hypotheses that stipulate that the impact of the shock or change (be it the introduction of minimum wage or the COVID-19 shock) on the outcome variable of interest (employment or leadership behavior) differs according to the relative degree or intensity in which the shock or change impacted on the outcome variable. And referring to our multi-level model (see eq. (6)): a similar specification in a two period before/after setting (see p. 235, Angrist & Pischke, 2009) would have been to have a regression where instead of the level of leadership behavior, the dependent variable would be the *change* in leadership behavior as dependent variable on the moderators like WFH potential, the crisis dummy (March 1st 2020) would then drop out of the model.

[4] See the summary of evidence and research in for instance https://voxeu.org/article/economic-uncertainty-wake-covid-19-pandemic which makes it clear that neither financial markets nor organizations as such saw the pandemic coming even right up until March 2020. See also: https://voxeu.org/article/financial-markets-and-news-about-coronavirus.

[5] https://ourworldindata.org/grapher/covid-stringency-index.

[6] https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

[7] Note that, as one reviewer pointed out: "At later periods of the COVID-19 crisis, i.e. post June 2020, the decision to WFH or adopt social-distancing measures or not consider COVID-19 as a serious threat is endogenously determined by e.g. the political context (say Texas versus New York City), scientific attitude towards appropriate policy against COVID-19 (say Sweden versus Switzerland), and public perception about COVID-19 measures and its impact on the economy (say Germany versus India). Therefore, for this period the possibility of correlation between error terms in a particular period and over the time period is present, which may result in inefficient estimates of the coefficients".

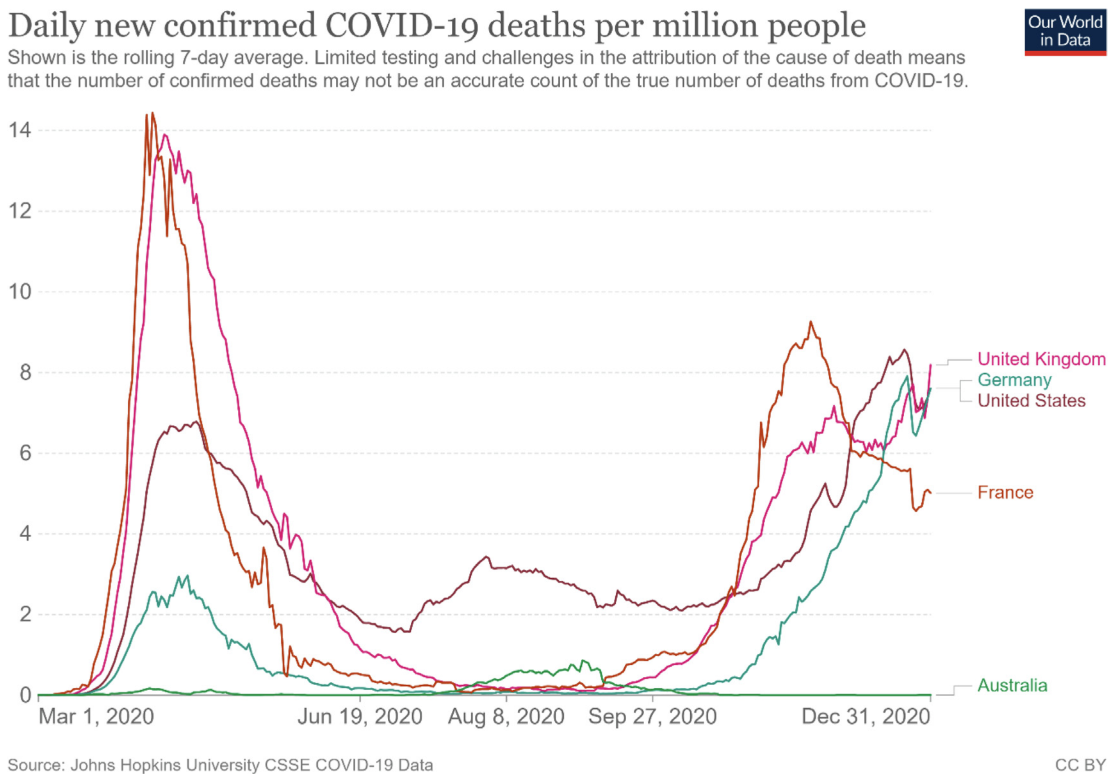


Fig. 1. COVID-19 deaths, March 1st 2020–December 31st (for five selected countries). https://ourworldindata.org/coronavirus

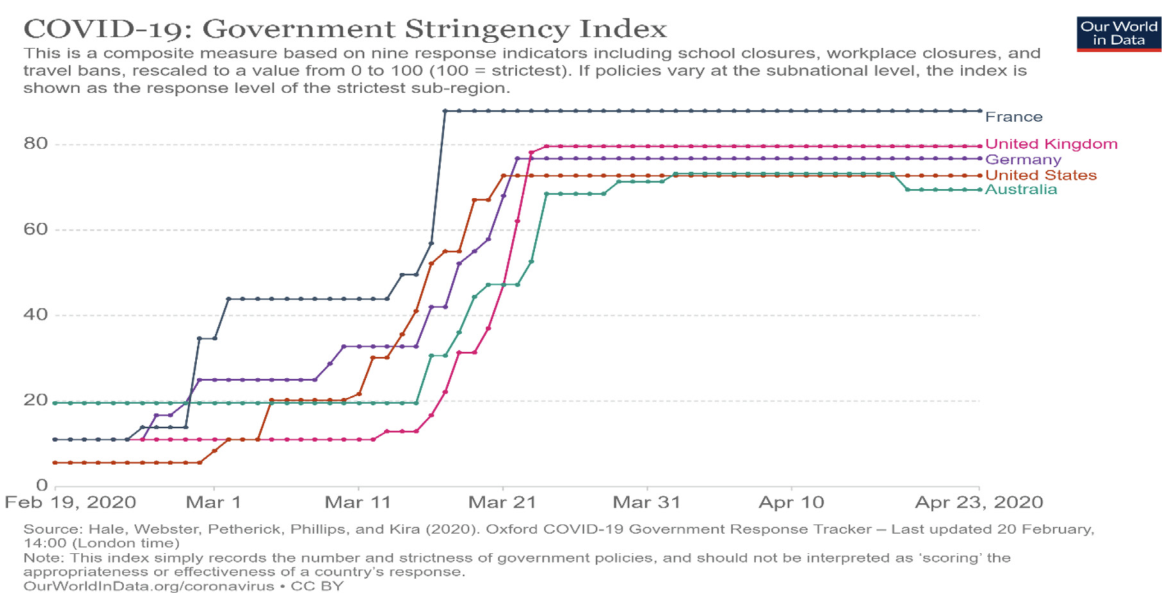


Fig. 2. Government Stringency Index, February 19th–April 23rd 2020 (for five selected countries). Source: Our World in Data.org. Both Fig. 1 and Fig. 2 construed by the authors using data options from https://ourworldindata.org/coronavirus.

large sample of countries (see for details: Euwema et al., 2007; Stoker et al., 2019; Van Emmerik et al., 2010).

For our present purposes, we accessed the relevant monthly assessment data for the period January 2019 to December 2020. We only included the assessment scores for each manager by their subordinates, so self-scores by managers were excluded. On average, each manager was assessed by five subordinates. To measure leadership behavior on the level of the individual manager, we took the average score from their respective subordinates. In total, we use a sample of N = 28,542 managers in our estimations for the full-sample period January 2019 to December 2020. The data, depending on the subsamples and model specifications used cover (at max) 619 organizations and 48 countries. We only include countries with at least 50 managers in our sample (see Appendix A).

The number of managers that participated in a KF program in both 2019 and 2020 is bound to be close to zero, and since the number of organizations that participated in the KF training programs in both years is limited as well, our main comparisons between the 'before' and 'after' shock leadership behavior will be done by comparing average effects. Since the number as well as identity of the managers varies per month in our data and a only a very small group of managers in our sample period took part in the KF program more than once, we are not dealing with a panel data set but with repeated cross-sections on a monthly basis. In addition, the number (and size of) organizations and hence the number of managers that took part in the KF program does vary per month. It is for these reasons that we focus only on the comparison of the average leadership behavior before and after March 2020 in a two-period setting.

Moreover, since the data availability per month varies quite considerably and also the main period of interest, the 1st lockdown period, is made up of 3 months only, to test for a common trend in the various treatment intensity variables, as is common in DID models, in the pre-treatment stage (in casu, in the sample period before March 1st, 2020) is not an option. As an alternative to a common trend test, one could also employ matching, to make sure that groups of managers before and after March 1st 2020 are as much alike as possible (except of course for the treatment itself). But, as we show in Appendix D, the pre- and post-March 1st samples are, reassuringly, already very much alike in both sectoral and individual manager (like age, gender, nativity) terms, such that matching would not yield a different pre-shock sample of managers in our view.

We are investigating the possible impact of an exogenous shock on leadership behavior, so we checked whether the participation in the programs of KF by organizations systematically differs 'before' or 'after' the arrival of the COVID-19 pandemic. Most of the samples by industry are more or less equal before and after March 1st. In two sectors, namely manufacturing and banking, we found that the manufacturing sample decreased, and that the banking sector significantly grew after March 1st. These effects were caused by the participation of two specific organizations. Therefore, these two organizations are not included in the samples we use for our estimations (see also the data Appendix D, with more information on the total sample composition per sector before and after March 1st 2020).

Moreover, on a country-level we find that the main (and perhaps not unexpected) difference is that the enrollment of Asian (= mainly Chinese) firms decreased relatively quite strongly during the first lockdown (March until June) for it to increase again in the second half of 2020. As a robustness check, we have therefore estimated our models with and without Chinese firms, but the results do not differ. We present the results with managers from Chinese firms included.

Measures

Directive leadership: Directive leadership was originally measured with 6 items ($\alpha = 0.83$). However, a CFA using all six items for directive (and five for participative) did not have a good fit: RMSEA > 0.1 and CFA, TLI < 0.9. We therefore dropped one item from the KF directive leadership scale. The results were satisfactory, with a RMSEA of 0.064 and CFI, TLI of both > 0.9 (see appendix D). The five-item scale is in line with another publication about the KF-dataset that uses directive leadership (Riisla, Wendt, Babalola, & Euwema, 2021); see Appendix B for the list of the items. All items used Likert-type scales, with answers ranging from 1 (non-directive) to 6 (very directive), with alternative answers on the extreme poles. Respondents were asked to "select the rating which best reflects your experience of < name manager>'s behavior". The scores of on average five subordinates were aggregated. We examined the justification for aggregating subordinates' responses by calculating the ICC (1) value, which was 0.38 (SE 0.0017) and can be considered as adequate (James, 1982) and/or not problematic (Bliese, Maltarich, & Hendricks, 2018). The mean $R_{wg}(j)$, defined as the within-team agreement score, for directive leadership is 0.49, which is a satisfactory level of inter-rater agreement.[8] As a robustness check, we also ran our analyses for groups with $R_{wg}(j) > 0.7$ for directive as well as participative leadership.

Participative leadership: Following Somech (2005, 2006) we define participative leadership as delegation of responsibilities and shared influence in decision making. Participative leadership was measured with 5 items ($\alpha = 0.83$), see Appendix B for the full list of the items. The set-up of the participative leadership scale was comparable to the directive leadership scale ranging from 1 to 6. Again, the scores of on average five subordinates were aggregated (ICC = 0.22, SE 0.0016). The mean $R_{wg}(j)$ for participative leadership is 0.60, which can be considered as adequate. In order to check whether directive and participative leadership are two different scales, we ran a two-factor and one-factor model comparison in CFA. Results clearly show that the two-factor solution is better compared to the one factor solution, according to the BIC: 56,745 versus 333,712 (see Appendix D).

Exogenous shock: Our independent variable is the variable that will differentiate our sample into a 'before' and 'after' sample, following the arrival of the COVID-19 crisis. It is impossible for our sample of multiple countries to pinpoint the shock to one single day, but from a monthly perspective it is clear that March 2020 was the month that the COVID-19 crisis took almost all countries across the globe into the 'pandemic' regime. As we explained in section 3, we consider the period March until June 2020 (the first lockdown period) as the most appropriate period to consider the COVID-19 crisis as an exogenous shock. As we have argued at some length in Section 3, summary statistics on the number of COVID infections and casualties as well as the stringency of policy measures taken indicate that March 2020 is the month of transition from pre- to post-shock. We will thus use *March 1st 2020* as a dummy variable to demarcate the before and after shock period.

Magnitude of the crisis: To test for the magnitude of the crisis, we include the number of COVID-deaths. Based on the data collected under the heading of Our World in Data, which provides a very rich and continuously updated source to gather information on various aspects of the current pandemic, we include country-data on the number of *COVID-deaths (as share of the population)* from March 1st 2020 onwards.

Power distance: We use the scores of Hofstede's cultural dimensions for each country (Hofstede, 2001; Hofstede & McCrae, 2004) to measure power distance, following Stoker et al. (2019). This indicator measures the degree of (in)equality between people within each country's society on a scale from 1 to 100. High values indicate that in a society people are very deferential to figures of authority and generally accept an unequal distribution of power. Following Beugelsdijk, Maseland, and Van Hoorn (2015), we can assume that scores for power distance are generally stable over time, and certainly for our two-year time frame.

WFH Potential (WFHP): Based on work by Dingel and Neiman (2020) on the sensitivity of sectors for WFH in general, we were able to establish different levels of WFH potential. Dingel and Neiman (2020) research for the USA, which has by now been backed up by similar WFH research for other countries[9], cannot only be used to rank sectors in terms of WFH potential, but also shows that this potential is tightly linked with the actual WFH practices from March 2020 onwards. We used their sector classification, see Appendix C, so as to group our KF sectors into low, medium and high WFHP sectors, denoted as the *WFHP*

[8] At the same time, and as suggested by one of the reviewers, we also conducted a within and between analysis (WABA) that suggests that individual (employee) effects are more likely than group level effects (at the level of the manager). Given that for our study, we are not interested in dyadic relationships between employees and their managers, by averaging the employeee-level scores we are assuming similar views about the manager and hence average out variation at the subordinate level to deal with the manager level only. See Van Mierlo, Vermunt, and Rutte (2009) as how to use WABA with RCM (we thank Heleen van Mierlo for her feedback on this matter).

[9] see https://www.economicsobservatory.com/what-has-coronavirus-taught-us-about-working-home.

(Low-Mid-High) variable in our data and estimations. Sectors were classified as low, medium or high potential sectors when the % of jobs that could be done at home was < 25%. 25–50%, or > 50% respectively.

Level of management: The KF data include a manager's level within the hierarchy of their organization, indicated by six categories that range from *Early-level individual contributor* to *Senior management*. We condensed the four lowest levels into a single category, to end up with three levels of management: *Low*, *Mid*, and *High*.

Control Variables: Given the fact that our sample consists of different managers in 2019 and 2020, we include a number of variables in order to control for individual differences between these managers. We control for gender, age and nativeness. Managers' *gender* is measured by including a dummy that equals 0 for male managers and 1 for female managers. We also include a dummy that codes whether the manager is a *non-native* (coded 0) or *native* (coded 1) of the country in which the organization resides. The *age* of managers is measured by their date of birth. We did not include tenure and the educational level of the manager because their inclusion would restrict our sample too much because of missing observations. Note that tenure is highly correlated with age, and education is not a very relevant control when it concerns the relationship between shocks and leadership behavior (see Stoker et al., 2019). We also added a control variable at the country level, *GDP per capita,* to control for the effect that the level of leadership behavior may also a function of the overall level economic development of a country in terms of human and technical resources.

Multi-level model specification

The model that we are going to estimate has a similar set up as the multi-level model that was used in Stoker et al. (2019). The basic idea is to arrive at a multilevel linear model that incorporates three levels of information. Level 1 includes variables pertaining to the individual manager, as well as the dummy variable *March 1st 2020* that takes the value 0 before and 1 after March 1st 2020 in our main estimations:

Level 1:

$$\begin{aligned} Leadership_{ijk} = {} & b_{oj|k} + b_{1j|k}(March1st2020) + \beta_2 Age_{ijk} \\ & + \beta_3 (LevelofManagement)_{ijk} + \beta_4 Native_{ijk} \\ & + \beta_5 Gender_{ijk} + \varepsilon_{ijk} \end{aligned} \tag{1}$$

This equation implies that the directive (or participative) leadership score for manager *i,* in an organization in sector *j,* located in country *k,* is a linear combination of the *March 1st 2020* dummy and the individual manager level control variables. In addition, the unobserved regression coefficients of the organization specific intercept and the *March 1st 2020* slope ($b_{0j|k}$ and $b_{1j|k}$, respectively) depend on other fixed and random effects, as shown in the Level 2 model below:

Level 2:

$$b_{0j|k} = b_{0k} + \beta_6 Sector_{jk} + \mu_{0j|k} \tag{2}$$

$$b_{1j|k} = b_{1k} + \beta_8 Sector_{jk} + \mu_{1j|k} \tag{3}$$

The Level 2 model implies that the organization *j* specific intercept in country *k* ($b_{0j|k}$) depends on the intercept specific to country *k* (b_{0k}), a sector specific covariate that codes whether the sector to which the organization belongs is either a low, medium or a high *WFHP* sector, and a random effect associated with that organization ($u_{0j|k}$). The organization specific slope effect, again, depends on a country specific time effect, the industry dummy and a random effect for the organization specific slope.

Level 3:

$$b_{0k} = \beta_0 + \beta_7 Country_k + \mu_{0k} \tag{4}$$

$$b_{1k} = \beta_1 + \beta_9 Country_k \tag{5}$$

Level 3 implies that the country specific slope effect depends on the overall mean and the country score, as well as a country random effect. *Country* is operationalized as either the score on the number of COVID-19 deaths (as percentage of the population), the score on Power distance, and as a control variable only, GDP per capita. Similarly, the random slope effect at the country-level depends on the overall time effect and country level covariate. We do also include a random country slope effect. The country-level variables were centered using the country level means and standard deviations, see also Antonakis, Bastardoz, and Rönkkö (2021). Substituting Level 3 into Level 2 and this in turn into Level 1 gives us the multilevel model to be estimated:

Table 1
Descriptive statistics (N = 28,542).

Variable	Mean/%	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
March 1st	24.2%	0.428	0	0	0	1
Female	28.7%	0.452	0	0	1	1
Age (in decades)	3.908	0.834	1	3	4	6
Native	88.2%	0.322	0	1	1	1
GDP per capita ($$)	33,014.030	22,991.510	1,876.525	10,286.580	51,404.430	80,504.400
Power distance	57.814	19.256	11	40	77	104
Deaths_pop	0.034	0.085	0	0	0.001	0.051
Participative leadership	4.570	0.609	1.000	4.200	5.000	6.000
Directive leadership	3.478	0.745	1.000	2.933	4.000	6.000
Management low	32.3%	0.468	0	0	1	1
Management mid	32.7%	0.469	0	0	1	1
Management_high	35.1%	0.477	0	0	1	1
WFHP high	27.3%	0.445	0	0	1	1
WFHP low	59.4%	0.491	0	0	1	1
WFHP mid	13.4%	0.340	0	0	0	1
March 1st:Deaths_pop	0.025	0.073	0	0	0	0
March 1st:WFHP low	0.161	0.368	0	0	0	1
March 1st:WFHP mid	0.055	0.229	0	0	0	1
March 1st: Management low	0.064	0.245	0	0	0	1
March 1st:Management mid	0.079	0.270	0	0	0	1
March 1st:Management high	0.098	0.298	0	0	0	1
March 1st:Power distance	14.617	27.385	0	0	0	104

Table 2
Correlations.

	March 1st	Female	Age	Native	GDP	Power distance	deaths_pop	Participative leadership	Directive leadership	Management low	Management mid	Management high
March 1st	1											
Female	−0.04	1										
Age	−0.02	−0.08	1									
Native	0.03	−0.02	−0.02	1								
GDP	−0.12	0.1	0.17	−0.19	1							
Power distance	0.08	−0.07	−0.17	0.13	−0.83	1						
deaths_pop	−0.08	0.14	0.09	−0.17	0.57	−0.49	1					
Participative leadership	0.07	0.04	−0.05	0.02	−0.04	0.04	0.02	1				
Directive leadership	0.05	−0.05	−0.11	0.1	−0.42	0.43	−0.3	−0.15	1			
Management low	−0.07	0.1	−0.21	−0.01	0.12	−0.1	0.13	0.02	−0.05	1		
Management mid	0	−0.04	0.02	0.01	−0.01	0.03	−0.03	−0.02	0.02	−0.48	1	
Management high	0.07	−0.06	0.18	0	−0.1	0.07	−0.1	0.01	0.03	−0.51	−0.51	1
WFHP high	−0.06	0.08	0	−0.1	0.2	−0.1	0.12	0.01	−0.13	−0.04	−0.01	0.05
WFHP low	0.09	−0.09	−0.05	0.05	−0.18	0.13	−0.16	−0.01	0.12	0.08	0.01	−0.09
WFHP mid	−0.05	0.02	0.07	0.06	0	−0.06	0.07	0	0	−0.06	0	0.06
March 1st:deaths_pop	0.61	0.02	0.02	−0.04	0.09	−0.12	0.33	0.04	−0.07	0	−0.01	0.01
March 1st:WFHP low	0.78	−0.05	−0.06	0.05	−0.16	0.12	−0.12	0.05	0.08	−0.03	−0.01	0.03
March 1st:WFHP mid	0.43	0	0.02	−0.04	0.06	−0.04	0	0.02	−0.04	−0.04	0.01	0.04
March 1st:Management low	0.46	0.02	−0.11	0.01	−0.02	0.01	0.01	0.04	0.01	0.38	−0.18	−0.19
March 1st:Management mid	0.52	−0.04	0	0.03	−0.06	0.04	−0.05	0.03	0.03	−0.2	0.42	−0.22
March 1st:Management high	0.58	−0.04	0.05	0.01	−0.1	0.06	−0.08	0.04	0.03	−0.23	−0.23	0.45
March 1st:Power distance	0.95	−0.05	−0.06	0.05	−0.23	0.22	−0.14	0.08	0.11	−0.07	0	0.07

	WFHP high	WFHP low	WFHP mid	March 1st:deaths_pop	March 1st:WFHP low	March 1st:WFHP mid	March 1st:Management low	March 1st:Management mid	March 1st:Management high	March 1st:Power distance
WFHP high	1									
WFHP low	−0.74	1								
WFHP mid	−0.24	−0.47	1							
March 1st:deaths_pop	0	−0.02	0.04	1						
March 1st:WFHP low	−0.27	0.36	−0.17	0.37	1					
March 1st:WFHP mid	0.4	−0.29	−0.1	0.32	−0.11	1				
March 1st:Management low	−0.04	0.07	−0.06	0.36	0.41	0.17	1			
March 1st:Management mid	−0.02	0.04	−0.02	0.31	0.39	0.23	−0.08	1		
March 1st:Management high	−0.03	0.03	−0.01	0.3	0.43	0.26	−0.09	−0.1	1	
March 1st:Power distance	−0.08	0.11	−0.06	0.46	0.78	0.35	0.42	0.49	0.57	1

$$Leadership_{ijk} = \beta_0 + \beta_1(March\ 1st\ 2020) + \beta_2 Age_{ijk} + \beta_3(Level\ of\ Management)_{ijk} + \beta_4 Native_{ijk} + \beta_5 Gender_{ijk} + \beta_6 Sector_{jk} + \beta_7 Country_k + \beta_8 Sector_{jk} * (March\ 1st\ 2020) + \beta_9 Country_k * (March\ 1st\ 2020) + \beta_{10}(Level\ of\ Management)_{ijk} * (March\ 1st\ 2020) + \mu_{0j|k} + \mu_{1j|k} * (March\ 1st\ 2020) + \mu_{0k} + \varepsilon_{ijk} \quad (6)$$

In line with for instance Johns (2018), our multi-level model specification is an example where the context is deemed to be relevant for individual leadership behavior, and where the context is defined at higher levels of aggregation than the organization and manager. Effectively, we will thus test a model with the dependent variable (leadership behavior) at the individual (=manager) level and the *cross-level moderating variables* at the micro (=level of management), meso (=sectoral level) and macro (=country level) levels. A cross-level model might be a more appropriate term than the more routinely used term multi-level model, to which we adhere for the sake of convenience.

Estimation results

Descriptive statistics

Descriptive statistics for all variables are presented in Table 1. As can be seen in Table 1, the average for the March 1st 2020 dummy is about 0.24, meaning that most of the observations fall in the pre-shock period, January 2019 until March 1st 2020. We still have almost 7,000 managers for the period March to December 2020. On the level of individual managers, we see that native, male managers of 40+ years old constitute the bulk of our sample, and that the managers are evenly spread across the three levels of management. As to the variables at the sector or country level, we note that most organizations fall within the low and high WFHP categories, and also that the variable COVID19-deaths has a skewed distribution. In Appendix D, we show sample comparisons before and after March 1st 2020, for both total sample and also for sub-samples for the our three WFHP variables as well as the three level of management variables. The main conclusion is that by and large these samples are very similar.

Table 3A
Results of the regression analysis for directive leadership, March until June 2020.

	Directive leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.094***	0.096***	0.097***	0.061	0.183***	0.124***	0.071**
	(0.025)	(0.025)	(0.025)	(0.043)	(0.037)	(0.026)	(0.031)
Management low	0.058***	0.041***	0.041***	0.036**	0.040***	0.041***	0.041***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Management mid	0.038***	0.026*	0.026*	0.027*	0.025*	0.026*	0.027*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Female	0.036***	0.040***	0.040***	0.040***	0.040***	0.040***	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Age	−0.022***	−0.026***	−0.026***	−0.025***	−0.025***	−0.025***	−0.025***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Native	−0.109***	−0.107***	−0.108***	−0.108***	−0.109***	−0.108***	−0.108***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
WFHP high		−0.070***	−0.070***	−0.070***	−0.059***	−0.069***	−0.071***
		(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)
WFHP mid		−0.055***	−0.054***	−0.054***	−0.054***	−0.053***	−0.054***
		(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
GDP			−0.175	−0.176	−0.176	−0.175	−0.175
			(0.237)	(0.237)	(0.237)	(0.237)	(0.237)
Power distance			0.173**	0.173**	0.173**	0.174**	0.175**
			(0.079)	(0.079)	(0.079)	(0.079)	(0.079)
Deaths_pop			−0.128	−0.128	−0.130	−0.133	−0.128
			(0.105)	(0.105)	(0.105)	(0.105)	(0.105)
March 1st:management low				0.106*			
				(0.059)			
March 1st:management mid				−0.015			
				(0.062)			
March 1st:WFHP high					−0.197***		
					(0.052)		
March 1st:WFHP mid					−0.001		
					(0.083)		
March 1st:deaths_pop						0.131***	
						(0.043)	
March 1st:Power distance							−0.060
							(0.045)
Constant	3.412***	3.461***	5.271**	5.289**	5.282**	5.275**	5.278**
	(0.070)	(0.071)	(2.513)	(2.510)	(2.515)	(2.512)	(2.514)
Observations	14,591	14,591	14,591	14,591	14,591	14,591	14,591
Log Likelihood	−14,145.380	−14,135.010	−14,131.670	−14,133.080	−14,127.470	−14,129.170	−14,132.960
Akaike Inf. Crit.	28,308.760	28,292.020	28,291.340	28,298.170	28,286.940	28,288.350	28,295.910
Random effects variance(τ_{00})	0.11	0.11	0.07	0.07	0.07	0.07	0.07

Note: N countries = 29. All estimations were performed in the R language (R Core Team, 2015) using packages NLME (Pinheiro, Bates, DebRoy, Sarkar, & Core Team, 2015) and STARGAZER (Hlavac, 2015).

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

Table 2 presents correlations among the variables used in our analysis.

As Table 2 shows, several of our control variables are correlated with directive (and participative) leadership. Note that we do not report significance because the correlation table includes different level variables. Below, we present estimation results for two sample periods: (1) the first lockdown period (March until June 2020) and, (2) the period March to December 2020. In these estimations, the main aim is always to compare before and after March 1st 2020. In Appendix D, we also show the correlation table at different levels (see also Eckardt, Crocker, & Tsai, 2020), including significance of the correlations.

For both samples, we will show the estimation results for directive leadership behavior as the dependent variable. As a robustness check, we will also estimate our models for participative leadership behavior. In keeping with the multi-level model specification, the lay-out of the estimation tables is identical by first checking if there is a significant change in leadership behavior when comparing 'before' and 'after' March 1st 2020 (including controls at the individual level), after which, and in line with our DID design, we will subsequently add the moderating variables at the organizational-, sector- and country-level and the relevant interaction effects.

All models shown below were estimated using random effects. For the estimations with fixed effects, the estimates are identical or within 1 standard error for the main variables of interest (available upon request). So, we opted for random effects, because it allows for multi-level interactions.

Estimation results for March until June 2020: The first lockdown period

Table 3A shows the estimation results for the first lockdown period, March 1st, until June 2020, for directive leadership, our main dependent variable of interest. In discussing these estimation results, we focus on the impact of the 'March 1st 2020′ dummy both as a direct determinant of directive leadership behavior and via possible interaction effects. This leaves us with a total of 14,591 managers, with 1,734 managers for the March until June 2020 period. Table 3A shows the estimation results for directive leadership. Row 1 indicates that directive leadership behavior on average did increase significantly after March 1st 2020 which is in line with hypothesis 1. Given the fact that the COVID-19 shock potentially affected all organizations and their managers, this result has to be qualified. The lack of a clear counterfactual that signals what leadership behavior would have been shown without the COVID-19 shock after March 1st 2020, or, in other words the lack of a control group, means that our findings are consistent with hypothesis 1 but the evidence should be seen as confirming a positive association between the COVID-19 shock and the increase in directive leadership,and not so much as evidence of a causal effect.

Column 1 shows that there are also significant estimation results for most of the control variables at the level of the individual managers: managers in low and middle-management as well as older, female or non-native managers show significantly higher levels of directive leadership behavior. Similarly, for key variables at the sector or country level (columns 2 and 3), there is also a significant relationship with the dependent variable, e.g., managers with organizations in sector with a relatively low WFHP display more directive leadership. Like Stoker et al. (2019), we also found a main effect for power distance (see column 3), indicating that in countries with a high power distance, the level of directive leadership is significantly higher. For the COVID-19 deaths, we do not find a significant main effect.

To test our moderating hypotheses 2 and 3a-c, columns (4)-(7) in Table 3A show the estimation results when the 'March 1st 2020′ dummy is interacted with the level of management, the WFHP variable, COVID-19 deaths, and power distance. The increase in directive leadership is found to be stronger after March 1st 2020 in countries that were hit harder by the COVID-19 pandemic, as measured by the number of COVID-19 deaths (see column 6, Table 3A), thereby supporting hypothesis 2. Fig. 3 illustrates that in countries that were less affected by the crisis (in terms of relatively fewer COVID-19 deaths per 100.000 inhabitants), directive leadership did not change as much after March 1st.

As for hypotheses 3a-c, we also find significant moderating effects for the first lockdown period. We do not find a significant interaction effect with power distance. We do, however, see significant interaction

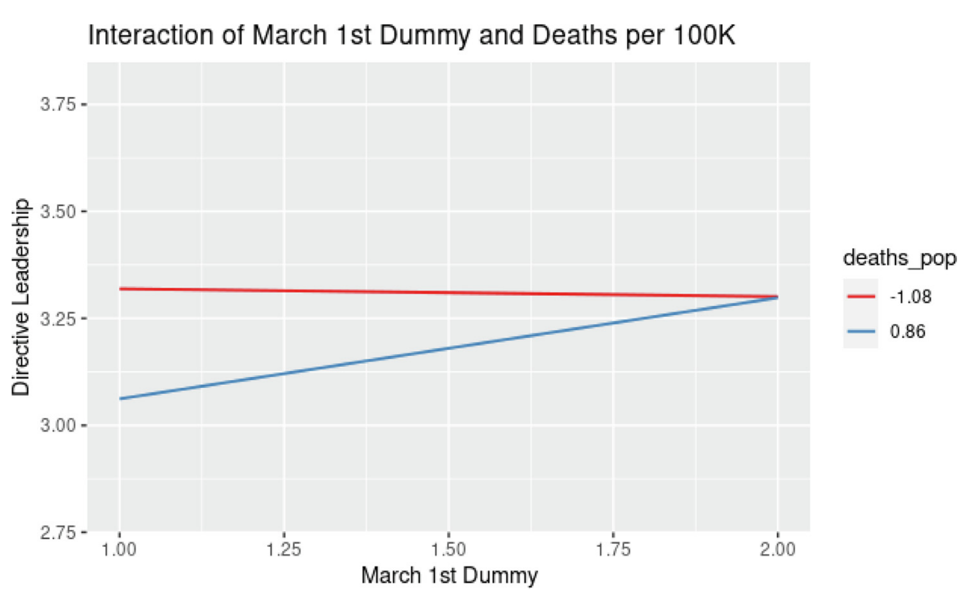


Fig. 3. Relationship between the COVID-19 crisis and the number of COVID-19 deaths per 100,000 inhabitants. Note: Deaths per 100,000 inhabitants is centered and −1.08 is the lowest value and 0.86 the highest value observed in the data. The figure shows the start and end point for the level of directive leadership for managers of the two countries with the highest (blue line) and lowest levels (red line) of COVID-19 deaths per 100 K. Start point refers to the avg. level of directive leadership before the COVID-19 shock, in casu January 2019-February 2020, and End point refers to the avg. level of directive leadership after the COVID-19 shock (after March 1st 2020), in casu for the period March-May 2020. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

effect for levels of WFHP (see Table 3A, column 5) in line with hypothesis 3b. To qualify this effect, Fig. 4 pictures the interaction results for the three levels of WFHP.

Fig. 4 shows two interesting results. First, we see that the 'starting' or steady level of directive leadership before March 1st is lowest in sectors where levels of WFHP are high, and that directive leadership in sectors with low levels of WFHP is highest. Second, the results show that after March 1st, there was an increase in directive leadership in sectors with low and mid levels of WFHP, but *not* in sectors where the WFHP is high. These results support hypothesis 3b, where we expected a relatively higher level of directive leadership under conditions of low levels of WFHP than under conditions of high levels of WFHP.

Finally, in line with hypothesis 3c, we do see a significant effect (at 10% level) of the level of management interacted with the crisis (see column 4, Table 3A). The results show that low level managers increased their directive leadership relatively more than managers at the middle or high levels. To qualify this effect, Fig. 5 plots this interaction effect of management level and the March 1st dummy.

In line with literature on directive leadership behavior and management levels, Fig. 5 confirms that directive leadership before March 1st is highest for low levels of management, although the differences with middle and high management are small. After March 1st, the increase in directive leadership is especially strong for low level managers. These results support hypothesis 3c, where we expected a stronger positive effect on directive leadership for low compared to high levels of management.

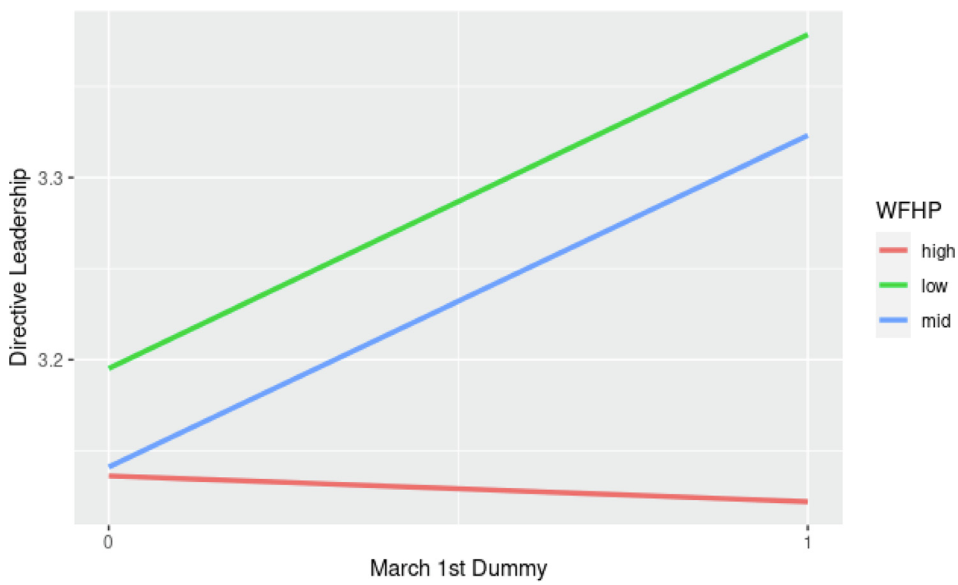


Fig. 4. Relationship between the COVID-19 crisis and directive leadership for low, middle and high WFHP.

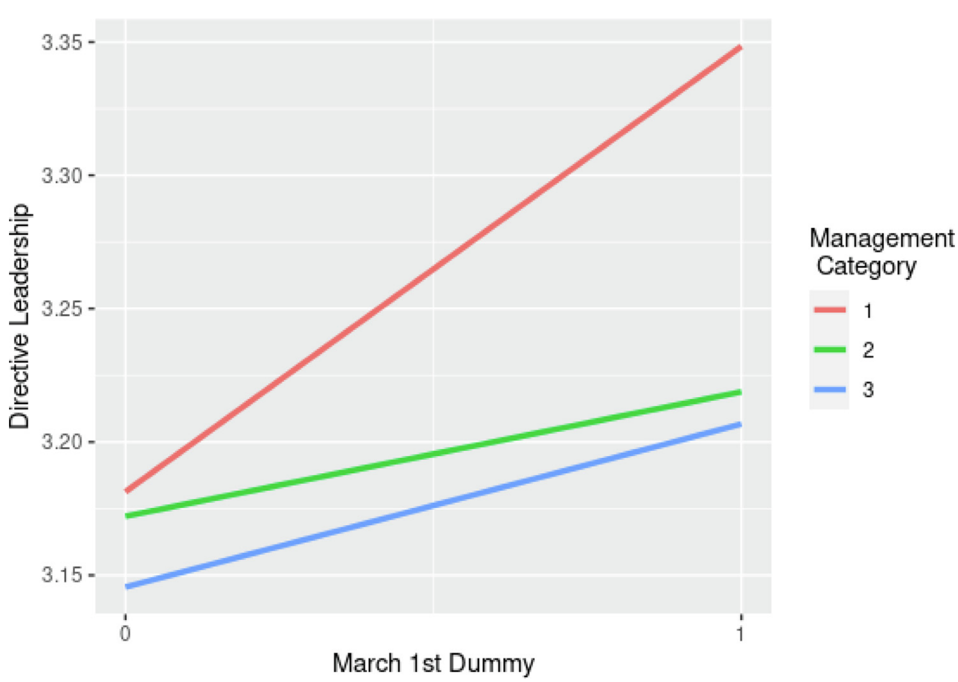


Fig. 5. Relationship between the COVID-19 crisis and directive leadership for low (1), middle (2) and high (3) levels of management.

A final and quite important remark concerns the causal interpretation of the results for hypotheses 2-3abc. In as far as we find significant results for the 'March 1st' dummy with COVID19-deaths, WFHP or the level of management, we do think that we are on a firmer footing, compared to the results supporting hypothesis 1, to give these results a causal interpretation. The reason is, and recall also our discussion of the DID research design, that in our DID set-up these variables are treatment intensity variables, which enable a test as to whether managers that belong to different treatment intensity categories do indeed on average show significant different change in directive leadership, when comparing before and after March 1st 2020. We obviously still lack a clear cut demarcation between a treatment and a 'formal' control group, but we can thus group managers in terms of relative treatment intensity along the lines of hypotheses 2 and 3abc.

Robustness checks

As a robustness check, we also ran our model for participative leadership, again in the first lockdown period (from March until June 2020). Results are presented in Table 3B.

From Table 3B we learn that, in line with the threat-rigidity hypothesis, there is no clear pattern for the possible effect of the COVID-19 crisis on participative leadership. We find no change and, if anything, a decrease of participative leadership in this timeframe (notably columns 5 and 6), but the result is not very robust. Like with directive leadership, there are a number of significant direct effects for our control variables. That is, female, older and native managers show more participative leadership. Moreover, we also see that managers in sectors with high WFHP display more participative leadership in this timeframe. With respect to the possible effects of our moderators, Table 3B shows one interesting result, namely with respect to the interaction between the March 1st dummy and WFH potential. Column 5 shows that managers working in sectors with both high- and mid-levels of WFHP, displayed relatively higher levels of participative leadership after March 1st. In Appendix D, sub f, we show two bar plots with confidence intervals for this estimation period for the interaction between the COVID dummy and WFHP for both directive and participative leadership, where the former is thus the bar plot version of Fig. 4.

As a second robustness check, we limited our sample based on the Rwg scores on directive and participative leadership. We reported that the mean within-team agreement rate $R_{wg}(j)$ was 0.49 and 0.60 for directive and participative leadership respectively. Our focus is not on

Table 3B
Results of the regression analysis for participative leadership, March until June 2020.

	Participative Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	−0.033	−0.038*	−0.038*	−0.025	−0.103***	−0.050**	−0.036
	(0.022)	(0.022)	(0.022)	(0.039)	(0.033)	(0.024)	(0.028)
Management low	0.012	0.023*	0.023*	0.026**	0.023*	0.023*	0.023*
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Management mid	−0.018	−0.011	−0.011	−0.012	−0.010	−0.010	−0.011
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)
Female	0.058***	0.055***	0.055***	0.055***	0.056***	0.055***	0.055***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Age	0.011*	0.015***	0.015***	0.015***	0.015**	0.015***	0.015***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Native	0.032**	0.035**	0.035***	0.035***	0.037***	0.035***	0.035***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
WFHP high		0.066***	0.067***	0.067***	0.061***	0.066***	0.067***
		(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
WFHP mid		−0.012	−0.013	−0.013	−0.020	−0.013	−0.013
		(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
GDP			−0.135*	−0.134*	−0.135*	−0.135*	−0.135*
			(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
Power distance			−0.027	−0.027	−0.027	−0.027	−0.027
			(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
deaths_pop			0.058*	0.058*	0.059*	0.060*	0.058*
			(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
March 1st:Management low				−0.054			
				(0.053)			
March 1st:Management mid				0.025			
				(0.056)			
March 1st:WFHP high					0.110**		
					(0.047)		
March 1st:WFHP mid					0.155**		
					(0.075)		
March 1st:deaths_pop						−0.059	
						(0.038)	
March 1st:pdi							0.004
							(0.040)
Constant	4.434***	4.390***	5.850***	5.839***	5.851***	5.848***	5.849***
	(0.033)	(0.034)	(0.826)	(0.824)	(0.828)	(0.825)	(0.826)
Observations	14,596	14,596	14,596	14,596	14,596	14,596	14,596
Log Likelihood	−12,568.030	−12,553.430	−12,558.110	−12,561.110	−12,558.220	−12,559.260	−12,560.400
Akaike Inf. Crit.	25,154.070	25,128.860	25,144.230	25,154.220	25,148.450	25,148.520	25,150.790
Random effects variance(τ_{00})	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Note: N countries = 29.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

the level of individual employees, and their individual perceptions of their managers. Instead of focussing on the dyadic relationship between employees and managers, we are only interested in comparing the leadership behavior of managers. By averaging the answers of employees, we are indeed assuming comparable perceptions about the manager for his or her group of employees. In doing so, we average out the variation at the employee level. As a robustness check we therefore also ran our analyses for a sample groups with higher R_{wg} scores, namely > 0.7, for both directive and participative leadership. As a supplement to the results shown in Tables 3A and 3B, the estimation results of the corresponding tables in Appendix D show the results for these sub-samples with $R_{wg}(j) > 0.7$. Clearly, the main results do not change.

Next, we will show the results for the sample period March to December 2020. Recall the discussion under 'reserch design', where we explained that endogeneity becomes an issue in this sample, so the results are to be cautiously interpreted.

Estimation results for March to December 2020

Table 4A presents the estimation results for directive leadership for the full-sample period of January 2019 to December 2020. The total sample includes just over 28,000 managers in total, and 6700 managers for the months March to December 2020. Again, this model is estimated using random effects.

Row 1 in Table 4A indicates that directive leadership behavior on average did not change significantly after March 1st 2020. Thus, there is no systematic significant main effect for directive leadership when comparing the periods before and after March 1st 2020 when we take into account the full period of our sample. This result could be seen as a rejection of our hypothesis 1, but this assumes that the exogeneity of the shock is unproblematic after June 2020 – we will return to this result in the discussion-section. Like we saw in Table 3A, there are again significant and comparable estimation results for most of the control variables at the level of the individual managers: managers in top-management as well as younger or native managers show significantly lower levels of directive leadership behavior, whereas the opposite holds for female managers (see column 1). Similarly, for key variables at the sector or country level (columns 2 and 3), there is also a significant relationship with directive leadership, e.g., managers working in organizations in a sector with a relatively high WFHP, display less directive leadership. Like Stoker et al. (2019), we found a main effect for power distance (see column 4), indicating that in countries with a high power distance, the level of directive leadership is, again, significantly higher.

Table 4A
Full sample results for directive leadership.

	Directive leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.024	0.014	−0.026	−0.012	−0.024	−0.028	−0.012
	(0.045)	(0.045)	(0.047)	(0.073)	(0.058)	(0.048)	(0.047)
Management low	0.132***	0.086*	0.089*	0.101*	0.089*	0.089*	0.093*
	(0.050)	(0.050)	(0.050)	(0.056)	(0.050)	(0.050)	(0.050)
Management mid	0.146***	0.113**	0.113**	0.112**	0.113**	0.113**	0.116**
	(0.047)	(0.047)	(0.047)	(0.055)	(0.047)	(0.047)	(0.047)
Female	0.097**	0.115***	0.115***	0.115***	0.115***	0.115***	0.114***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Age	−0.126***	−0.140***	−0.139***	−0.139***	−0.139***	−0.139***	−0.139***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Native	−0.209***	−0.218***	−0.218***	−0.218***	−0.218***	−0.218***	−0.217***
	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
WFHP high		−0.362***	−0.365***	−0.364***	−0.365***	−0.364***	−0.366***
		(0.049)	(0.049)	(0.049)	(0.055)	(0.049)	(0.049)
WFHP mid		−0.032	−0.034	−0.034	−0.029	−0.035	−0.040
		(0.061)	(0.061)	(0.061)	(0.067)	(0.061)	(0.061)
GDP			−0.739***	−0.739***	−0.742***	−0.731***	−0.744***
			(0.232)	(0.232)	(0.232)	(0.235)	(0.233)
Power distance			0.857***	0.857***	0.856***	0.863***	0.887***
			(0.250)	(0.251)	(0.251)	(0.252)	(0.251)
Deaths_pop			−0.367**	−0.367**	−0.366**	−0.369**	−0.368**
			(0.179)	(0.179)	(0.179)	(0.179)	(0.180)
March 1st:Management low				−0.057			
				(0.109)			
March 1st:Management mid				0.007			
				(0.105)			
March 1st:WFHP high					0.003		
					(0.107)		
March 1st:WFHP mid					−0.025		
					(0.145)		
March 1st:deaths_pop						0.008	
						(0.030)	
March 1st:Power distance							−0.194***
							(0.061)
Constant	17.848***	18.044***	25.306***	25.302***	25.328***	25.219***	25.344***
	(0.301)	(0.303)	(2.304)	(2.304)	(2.307)	(2.330)	(2.309)
Observations	28,542	28,542	28,542	28,542	28,542	28,542	28,542
Log Likelihood	−73,639.550	−73,615.330	−73,594.680	−73,597.220	−73,597.000	−73,597.220	−73,591.430
Akaike Inf. Crit.	147,297.100	147,252.700	147,217.400	147,226.400	147,226.000	147,224.400	147,212.900
Random effects variance(τ_{00})	0.12	0.12	0.0.05	0.05	0.05	0.05	0.05

Note: N countries = 48.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Columns (4)-(7) in Table 4A show the estimation results when the 'March 1st 2020′ dummy is interacted with the level of management, the WFHP variable, COVID-19 deaths, and power distance respectively, following the moderating effects as casted in hypotheses 2 and 3a-c. These columns show that there are no significant interaction effects between these variables and the March 1st dummy for directive leadership. So, in the full sample period, we do not find support for our moderating hypotheses. We do find a significant interaction of the crisis with power distance. However, the interaction-effect is negative, indicating that in countries with a high power distance there was a decrease in directive leadership after the crisis.

Table 4B shows the corresponding estimation results for participative leadership. Notably, we find a very clear and consistent significant change in participative leadership, because the level of participative leadership is on average significantly higher after March 1st 2020. This significant change of participative leadership also remains unchanged when adding the various moderating variables in our multi-level model from column (4) to column (7). We do not find significant interaction effects for the selected interaction variables.

Overall, the full sample results for the period March to December 2020 thus show two clear outcomes: on average no change in directive leadership behavior, and a significant increase of participative leadership behavior. When set against the estimation results for the first lockdown period (Tables 3A and 3B), the estimation results in Tables 4A and 4B show different results, especially where it concerns the main effect of the crisis on leadership.

Just like with the estimation results underlying Tables 3A and 3B for the 1st lockdown period, we re-ran our estimations by only including groups for which the within group agreement score was $R_{wg}(j) > 0.7$ for the COVID-19 period June-December 2020. The estimation results are similar to the results shown for this sample period in Tables 4A and 4B, and can be found in Appendix D sub c.

We also estimated the exact same models for directive and participative leadership, but now for the COVID-19 period June to December 2020 only, so excluding the first lockdown period. These results are in line with the findings in Tables 4A and 4B: we find no significant change or even a decrease for directive leadership behavior, and a significant increase in participative leadership (results are also available upon request).

Both the estimation results for March to December 2020, as well as the results for June to December 2020, must be seen in terms of mere associations between the crisis and leadership behavior, instead of in

Table 4B
Full sample results for participative leadership.

	Participative Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.075***	0.076***	0.073***	0.070***	0.058***	0.074***	0.073***
	(0.009)	(0.009)	(0.009)	(0.013)	(0.011)	(0.009)	(0.009)
Management low	0.014	0.020**	0.021**	0.020*	0.021**	0.020**	0.020**
	(0.009)	(0.009)	(0.009)	(0.011)	(0.009)	(0.009)	(0.009)
Management mid	−0.001	0.004	0.004	0.001	0.003	0.004	0.003
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)
Female	0.036***	0.034***	0.034***	0.034***	0.034***	0.034***	0.034***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Age	−0.003	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Native	0.038***	0.040***	0.040***	0.039***	0.040***	0.039***	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
WFHP high		0.050***	0.051***	0.050***	0.045***	0.051***	0.051***
		(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)
WFHP mid		−0.002	−0.002	−0.002	−0.019	−0.002	−0.002
		(0.011)	(0.011)	(0.011)	(0.013)	(0.011)	(0.011)
GDP			−0.054**	−0.054**	−0.053**	−0.055***	−0.054**
			(0.021)	(0.021)	(0.021)	(0.022)	(0.021)
Power distance			−0.028	−0.028	−0.028	−0.029	−0.029
			(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Deaths_pop			0.041***	0.041***	0.040***	0.042***	0.041***
			(0.014)	(0.014)	(0.014)	(0.015)	(0.014)
March 1st:Management low				0.001			
				(0.021)			
March 1st:Management mid				0.008			
				(0.020)			
March 1st:WFHP high					0.025		
					(0.020)		
Mrch 1st:WFHP mid					0.081***		
					(0.027)		
March 1st:Deaths_pop						−0.003	
						(0.006)	
March 1st:Power distance							0.007
							(0.011)
Constant	4.494***	4.467***	5.005***	5.006***	4.991***	5.015***	5.003***
	(0.027)	(0.028)	(0.214)	(0.214)	(0.213)	(0.214)	(0.213)
Observations	28,553	28,553	28,553	28,553	28,553	28,553	28,553
Log Likelihood	−25,844.370	−25,835.420	−25,838.300	−25,844.280	−25,839.280	−25,842.420	−25,841.670
Akaike Inf. Crit.	51,706.730	51,692.840	51,704.590	51,720.560	51,710.560	51,714.830	51,713.340
Random effects variance(τ_{00})	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Note: N countries = 48.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

terms of causality. Since we argue that endogeneity (increasingly) becomes an issue after June 2020, it is difficult to isolate the relevance of the COVID-19 shock for the findings for June-December 2020. In most countries, the first lockdown period thus ended in June 2020 and government measures became less stringent or restrictive. The latter is evident via a decrease of the Government Stringency Index, on which Fig. 2 is based, from June 2020 onwards in most countries. Also (see Fig. 1), the number of COVID-19 deaths and infections came down, and the predicted economic crisis with mass firm closures and lay-offs did not materialize.

This all suggests that the external threat that COVID-19 constitutes for managers and their organizations, may have waned off (at least temporarily) in the second half of 2020. At the same time, *actual* working from home that, ceteris paribus, possibly indeed stimulates managers to show higher (lower) levels of participative (directive) leadership behavior increased, underpinning our findings for the June-December 2020. Having said so, we can only speculate because, and in contrast with the first lockdown period, cause and effect explanations become difficult if the exogeneity of the COVID-19 shock and its impact on leadership behaviors gets more problematic to begin with, because the implications of the shock for managers are increasingly also the result of man-made interventions at the governmental or organizational level.

Discussion and conclusions

Conclusions

In this study, we investigate the effect of the COVID-19 crisis as an exogenous shock on leadership behavior across a wide range of organizations, sectors and countries. In line with the threat-rigidity hypothesis (Staw et al., 1981), we cast the unprecedented scale and nature of the shock as an external threat. Assuming that the COVID-19 shock in March 2020 is an exogenous shock, we present empirical evidence whether and how this shock affected leadership behavior. Based on the threat-rigidity hypothesis, we propose in hypothesis 1 that directive leadership behavior will have increased in the wake of the COVID-19 crisis. Additionally, we develop four hypotheses to explain how contextual variables moderate this relationship between the COVID-19 crisis and leadership behavior. We propose that the magnitude and well-learned responses strengthen the effect of the crisis on directive leadership (Staw et al., 1981).

The magnitude of the crisis effect is measured by COVID-19 deaths (in percentage of the population), whereas the well-learned responses are operationalized via power distance, the level of WFHP and the level of management. Following the threat-rigidity hypothesis, we are predominantly interested in the effect of the crisis on directive leadership. As a robustness check, we also investigate the effect on participative leadership (Stoker et al., 2019). Our research design is a DID-design with a treatment intensity variable (Angrist & Pischke, 2009), where our moderators are used as the treatment intensity variables. Our demarcation of the exogenous shock of COVID-19 is the first lockdown period, that is March 1st 2020 until June 2020, but we also ran our models for the period March 1st to December 2020, although exogeneity of the shock is problematic in this period.

Our estimation results are in line with hypothesis 1: we find an increase in directive leadership in the period March to June 2020, where this result is best seen as significant positive association between the COVID-19 shock and directive leadership. The lack of clear-cut counterfactual cautions against a causal interpretation. Moreover, we also find confirmation for almost all proposed moderators during this period. Given our research design of a DID model, where the moderators are to be seen as treatment intensity variables, we see however these results as providing evidence for a causal impact of the COVID-19 shock on directive leadership behavior. As hypothesized, in countries with relatively high numbers of COVID-19 deaths, we find a stronger increase in directive leadership. This finding indicates that the magnitude of the crisis matters for the effect of the shock.

The remaining hypotheses refer to the effect of well-learned responses. At the sectoral level, results confirm our hypothesis that in sectors with low- and mid-level of WFH potential, there is an increase in directive leadership. Regarding the level of management, we find that low level managers show a somewhat higher increase in directive leadership following the crisis, thereby also confirming our hypothesis. With respect to the possible interacting effect of power distance, our results do not confirm our hypothesis; we do not find a stronger effect of the COVID-19 crisis on directive leadership in countries with high power distance. By way of robustness check, and as further confirmation of the threat-rigidity hypothesis, our results for participative leadership show no significant change in this sample period.

For our second COVID-19 crisis period, which covers March to December 2020, we ran the same models. These results are different, and partly even opposite to our findings for the first lockdown period. Notably, we do not find a significant change in directive leadership, whereas we do find a significant increase in participative leadership. We do not find any meaningful significant interaction-effects of our moderating variables in line with our hypotheses.

When comparing the results for these two periods (March until June, versus March to December), we have previously argued that only the first period really qualifies as a period where the crisis can be seen as an exogenous shock. Crucially, the estimation results of this period can therefore be given a causal interpretation. However, for the period March to December 2020, this is no longer the case. In our view, for this period, the shock becomes at least partly endogenous; therefore, the estimation results for this period have to be regarded as merely associations between the crisis and leadership behaviors.

Nevertheless, our results for the time frame March to December 2020 are intriguing, because they might hint at two consequences of the COVID-19 crisis for leadership. First, and in line with earlier findings (Stoker et al., 2019), the results seem to suggest that the increase in directive leadership was temporary. After the initial increase in the period March to June 2020, we see that this increase in directive leadership stopped after June. Notably, see Fig. 1, this result coincides with the sharp drop in COVID-19 deaths, which we casted as the main indicator for the magnitude of the crisis. Possibly, we see here that the decrease of the threat, that is to say once the immediate threat of COVID-19 waned off, no longer went along with an increase in directive leadership.

Second, we find a significant increase in participative leadership in the timeframe March to December 2020. The context of WFH, which was relevant for a lot of the managers and employees in our sample, involved that managers had to lead from a distance. In such a context, managers are expected to exert a more participative style, by giving autonomy to followers and by delegating responsibilities (Contreras, Baykal, & Abid, 2020). Our results indicate that the increase in participative leadership might be driven by the different stages of the COVID-19 crisis in the course of 2020. It is conceivable that these behaviors became more routine for managers after a couple of months after the start of the COVID-19 crisis and the corresponding working from home situation (Bonet & Salvador, 2017). Clearly, this result is something for further research.

Theoretical contributions

As we stated in the introduction of our paper, our study contributes to leadership and management research in three ways. First, by investigating the effects of an exogenous shock on leadership behavior, we show how a global shock like the COVID-19 crisis can be used accordingly, and how the subsequent results can be given a causal interpretation via a DID research design (Antonakis et al., 2019; Sieweke & Santoni, 2020). Moreover, our contribution also concerns the fact that we have fine-grained (monthly) data which allow to be much more precise as to the time-frame of the exogenous shock.

Second, we are among the first to analyze how the current and still ongoing COVID-19 crisis impacts on leadership behavior and we thereby contribute to the growing literature on crises and specifically on the effect of the COVID-19 crisis on leadership. Recently, Rudolph et al. (2021) strongly advocated for such endeavors: "*such research could attempt to constructively replicate a recent study on the threat-rigidity hypothesis (….) it would be interesting to compare and explain the engagement in, and preference for, different leadership behaviors before and after COVID-19*" (Rudolph et al., 2021, p. 20).

As a third and final contribution, and by employing a DID research design setting, we are able to show how the effect of the threat of the COVID-19 pandemic is moderated by contextual variables. This provides further evidence on the relevance of the threat-rigidity hypothesis, in particular with respect to the magnitude and well-learned responses dimensions of this hypothesis. Hereby, we also replicate findings of Stoker et al. (2019). In addition, we not only confirm the threat rigidity hypothesis but our analysis extends the findings from Stoker et al. (2019). That is, we find that the level of management is a crucial contextual factor when it comes to the impact of a shock or crisis on directive leadership. We show that especially lower level managers are affected by such a shock, that is, show an increase in directive leadership. Our results clearly indicate that managerial level is an important variable when it comes to research on exogenous shocks and leadership.

Taken together, our findings can be thus seen as strong support for the threat-rigidity hypothesis (Staw et al., 1981, Stoker et al., 2019) when it concerns the period in the intermediate aftermath of the arrival of COVID-19, that is to say during the first lockdown period. In this period, our findings confirm that an external threat can lead to an increase in directive leadership. One obvious way to gather additional and more confirmative evidence on the effect of the COVID-19 on leadership behavior would be to collect more data as time moves on, so as to see how leadership behavior evolves in 2021 and beyond. This is certainly something one could explore further, but the trade-off here is that more data, and using different sub-periods, also too easily could imply that the endogeneity problem increases, because the impact of the COVID-19 crisis on organizations is increasingly the result of decisions made by organizations and managers themselves, thereby calling into question the research design that was the starting point of our article (and this Special Issue).

Limitations and future research directions

This special issue of *The Leadership Quarterly* has as a common denominator that exogenous shocks create the possibility to investigate the relevance and the impact of context on leadership in a rigorous manner. This rigor is not only a function of the shock being exogenous, thereby potentially allowing for causal inference, but also of the choice of an appropriate research design (Sieweke & Santoni, 2020; Antonakis, 2017; Garretsen et al., 2020). In this paper, we use the arrival of the COVID-19 crisis as an exogenous shock from March 2020 onwards, and we employ a specific differences-in-differences (DID) design to study the impact of the COVID-19 shock on leadership behavior. In doing so, we use the heterogeneity of contextual factors that allow us to capture the variation of this shock across countries, sectors and their organizations and managers. In our DID set-up, the contextual or moderating factors are used as treatment intensity variables.

We acknowledge that the exogeneity of the COVID-19 shock decreases over time. In particular, exogeneity is best guaranteed when focusing on the initial period of the shock, that is the period of the first lockdown (March until June 2020). Confirming the threat-rigidity hypothesis, we find evidence that directive leadership behavior increased in this period, that is the immediate wake of the COVID-19 shock. In the first lockdown period, the increase in directive leadership signals a rigidity reflex of managers. We also find evidence that contextual factors moderate the effect of the COVID-19 crisis on leadership. For our full crisis sample period (March to December 2020), we do not find a change in directive leadership, but instead the results show a significant increase in participative leadership. But as we explained at length in our paper, the exogeneity of the shock becomes questionable, and therefor also the alleged causality of the relationship between the COVID-19 crisis and leadership.

Our study has some other limitations. First of all, and despite the richness of our data set, one would ideally like some of the contextual factors, like WFHP, to be measured on the organizational level or, even better, on the level of the individual manager. Unfortunately, we do not have access to such data. Similarly, it would be interesting to have more personal information about the individual manager, next to the biographical variables (like gender and age) that we were able to include in our analyses. Specifically, the possible role of emotions of leaders seems relevant, because Madera and Smith (2009) show that evaluation of leadership in a crisis situation can be influenced by the leader's emotional reaction. Relatedly, when it comes to the assessment by employees of the leadership behaviour by their manager before and after the shock, there is considerable within-group variation, indicating that subordinates disagree in their assessment. Since we are, however, not interested in dyadic relationship between employees and their respective manager, but are instead focussing on drawing macro-level conclusions, we are to some extent assuming similar views about the manager thereby averaging out some variation at the subordinate level to deal with the manager level.

Third, the shock we are investigating is a truly global shock that, notwithstanding the varying impact across countries and sectors, potentially hit all organizations and their managers in our data set. From the perspective of causal inference, we are therefore limited to studying the difference between the before and after of the shock only. This precludes the use of a more standard DID-design with a treatment and control group. Having said so, we think that our DID design with a treatment intensity variable is a valid alternative. Stronger still, data limitations as well as, see also above, the rather tight time-window that the shock could be deemed to be exogenous, led us to estimate a two-period model. In such a two-period setting where we group together and average the leadership behavior before versus after the shock, a more stringent testing of causality that tests for common trends between relatively more and less 'treated' groups of managers is excluded. In such a two-period setting, claims about causality carry this disclaimer. In future research, and with more data becoming available as we move further away from the initial beginning of the COVID-19 shock, the deployment of a more stringently defined research design, DID or otherwise, would certainly be useful as a check on our findings. In a similar vein, future research could also try to include more detailed statistics on fixed and/or random effects estimations employed to solidify the respective estimations (see for instance Appendix C in Antonakis et al. (2021) for various codes as how to do so using STATA[10]).

Another limitation, or perhaps more appropriately, a suggestion for further research, is that at the time of the writing of the article, (the implications of) the COVID-19 crisis is very much still ongoing. Collecting more data as time goes by is thus a possible goal for further research, also since we have argued and shown that 'time' was a highly relevant factor in assessing the impact of the shock on leadership behavior in 2020. Extending the data set beyond 2020 would, hypothetically, allow for additional testing of our current or even new hypotheses. From a research design perspective however, doing so

[10] For a very useful review on this matter see https://www.princeton.edu/~otorres/Panel101.pdf, which provides a step-by-step estimation instructions for the case of STATA.

would put great extra weight on causal inference. Extending the data set for the crisis at hand creates also a drawback in the sense that the health, economic and also WFH impact of the COVID-19 crisis on leadership behavior will invariably become even more endogenous through interventions by not only policy-makers, but also by organizations and managers themselves.

Our study deals with the antecedents of (changes in) leadership behavior. A question for further research would also be to try to link the shock-induced changes in leadership behavior to organizational outcomes. Our current data set only offers outcome measures on the team level, but leaving aside this limitation, a stumbling block for this avenue of research would - again - be the endogeneity of such outcome measures. Further drawing upon the growing literature in leadership research (see for a review Sieweke & Santoni, 2020) that exploits or creates an experimental setting to infer causality from leadership (behavior) to outcome or performance measures, seems the next step to link the literature on the impact of shocks on leadership to their subsequent impact on performance (see also Garretsen et al., 2020).

Practical implications

Last but not least, the focus of the present study has strongly been on its relevance for leadership researchers. This academic focus is very much in keeping with the goals of the Special Issue. Having said so, we do believe that our study and its main findings are relevant for leadership practitioners. Understanding how large exogenous shocks like COVID-19 might impact on leadership behavior is a necessary condition or first step for organizations and their managers to try to deal with such shocks.

In addition, understanding why and how this impact might vary across different contexts is also important for leadership practice and HR-professionals. The results clearly show that the COVID-19 crisis especially led to an increase of directive leadership of lower level managers, and of managers who worked in sectors with low and mid levels of WFHP. However, whether such an increase in directive leadership in these contexts is effective is rather doubtful (see also Stoker et al., 2022), and such behavior might even be detrimental to performance or innovation (Somech, 2005). These results are especially relevant for HR-practitioners in organizations, who can monitor possible changes in leadership behaviors especially for these groups, and assist managers in becoming aware of the effect of a crisis on their behavior.

As we stated in the introduction, the current COVID-19 crisis is an opportunity for researchers, but it is obviously a curse for leaders and their organizations in the real world. Doing rigorous and novel research on shocks and leadership is one way how academic research may inform and help leadership practitioners to deal with the consequences of the COVID-19 pandemic.

Data availability

The authors do not have permission to share data.[1]

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[1] Requests with respect to the use of the dataset can be made to the authors, who will inform Korn Ferry about this request.

Appendix A. Allocation of organizations and managers across 48 countries and 32 sectors (sample Jan 2019–Dec 2020)

Country	Organizations	Managers
Argentina	15	61
Australia	64	916
Austria	19	59
Belgium	33	1231
Brazil	57	1075
Bulgaria	8	83
Canada	52	593
Chile	10	40
China	151	4053
Columbia	19	81
Costa Rica	5	49
Czech Republic	18	103
Denmark	16	106
Ecuador	4	162
Finland	10	197
France	55	467
Germany	67	808
Greece	10	57
Guatemala	3	145
Hong Kong	36	112
Hungary	14	26
India	56	2021
Indonesia	27	251
Ireland	22	282
Israel	8	121
Italy	34	216
Japan	79	2082
Malaysia	27	350
Mexico	30	209
Netherlands	46	501
New Zealand	13	123
Norway	10	67
Peru	11	319
Philippines	13	50
Poland	36	715
Portugal	14	59
Romania	13	141
Russia	23	114
Slovakia	11	201
South Korea	43	335
Spain	41	352
Sweden	19	163
Taiwan	21	128
Thailand	25	54
Turkey	29	925
United Kingdom	96	1507
United States	199	5158
Vietnam	20	112

Industry (Korn Ferry sector-classification)	Organizations	Managers
Manufacturing	55	2805
Food Products	33	2426
Consumer Products (excl. Food & Beverage)	22	714
Chemical & Related Products	25	428
Pharmaceuticals	46	2331
Technology	74	4243
Telecommunications	10	249

Appendix A (*continued*)

Country	Organizations	Managers
Financial Services	30	974
Banks/S&L's	23	1018
Insurance	21	1529
Health	28	564
Utilities	13	394
Construction	15	899
Diversified Conglomerates	11	840
Agriculture	7	149
Petroleum	12	787
Mining	5	94
Real Estate	25	745
Retail	23	475
Hospitality and Tourism	2	21
Entertainment/Recreation	6	99
Wholesale Trade	4	17
Transportation	14	398
Communications	6	117
Broadcast Media	2	199
Professional Services	46	2072
Legal	4	144
Professional Services – 3rd Parties	27	659
Education	20	463
Public Administration	8	310
State & Local	10	748
Associations	7	52

See also Appendix C as to how these 32 sectors and thus the associated organizations and managers are classified in terms of working from home (WFH) potential.

Appendix B. Items for directive and participative leadership

Directive Leadership Items
Requires employees to provide detailed updates
Expects employees to carry out instructions immediately
Quickly corrects team members that deviate from directions
Monitors what employees are doing very closely
Pays very close attention to what team members are doing

© Korn Ferry

Participative Leadership Items
Encourages the team to make decisions for themselves.
Prefers that decisions be made through consensus
Keeps everyone in the team involved and well-informed about organizational issues that may affect them.
Encourages employees to participate in most decision-making
Regularly adopts new ideas from the team.

© Korn Ferry

Appendix C. Working from home potential (WFHP) – sector classification

Agriculture	Low	Manufacturing	Low
Associations	High	Mining	Mid
Banks/S&Ls	High	Petroleum	Mid
Broadcast Media	High	Pharmaceuticals	Low
Chemical & Related Products	Low	Professional Services	High
Communications	High	Professional Services-3rd parties	High
Construction	Low	Public Administration	Mid
Consumer Products (excluding Food & Beverage)	Low	Real Estate	Mid
Diversified Conglomerates	Low	Retail	Low
Education	High	State & Local	Mid
Entertainment/Recreation	Mid	Technology	Low
Financial Services	High	Telecommunications	High
Food Products	Low	Transportation	Low
Health	Mid	Utilities	Mid
Hospitality and Tourism	Low	Wholesale trade	Low
Insurance	High		
Legal	High		

Source: Table 3 in Dingel and Neiman (2020) gives for 2 digit NAICS sectors the share of jobs that could be done from home (WFH). We used their (unweighted) sector shares by first allocating each sector code to each of the KF sectors mentioned in our table above and by using the following cut-off values for WFH shares (in terms of possibility to work from home in that sector): <25% = Low, 25–50% = Mid and >50% = High. See Appendix A for the allocation of organizations and managers in our sample over these sectors.

Appendix D

This data appendix groups together additional information on the sample used, and various alternative measures and specifications. It thereby provides background information on a number of topics from the main text.

(a) Sample composition across industries before and after March
(b) CFA analysis
(c) Two alternative tables for 3a and 3b, for 1st lockdown, with Rwg > 0.7
(d) Sample comparisons: I) Total sample before and after March 1st II) the same but now for 3 WFHP sub-samples and 3 level of management sub-samples
(e) Correlation table in levels
(f) Confidence intervals for bar plots

Appendix D, sub a: Sample composition

As to possible selection bias with respect to our sectors: we checked for the possible differences in participating organizations. Here, we do see a difference in participating organizations before and during COVID-19. This is the consequence of the way how the data are collected. If a program runs for organization X during a (short) period, you will not find this organization earlier or later in the database. Importantly, as long as this fluctuation is random (free of bias), it does not hurt our research design.

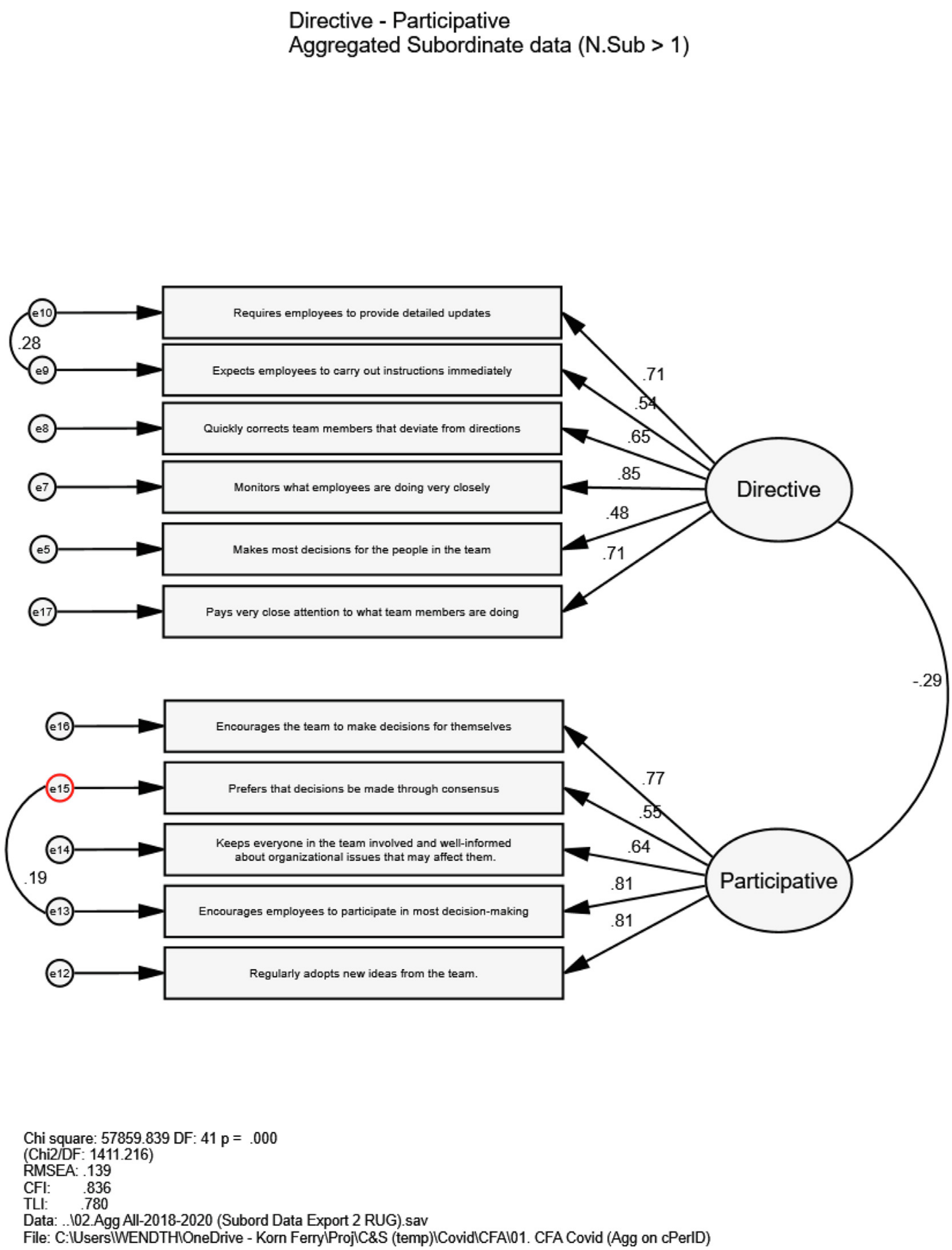
We therefore scanned the participation of organizations clustered by industry before and during COVID-19. Most of the samples by industry are more or less equal before and after March 1st (see table below). But given these results, we checked whether there might be an issue in two industries, namely the manufacturing industry and the banks (highlighted in Yellow below). We see that the Manufacturing sample shrinks, and that the Banking sector grows.

Looking at the specific firms in the Manufacturing industry, we can conclude that this drop can be explained by only one firm, who had 2.683 managers in 2019 till March 1st, and only 293 after March 1st. For the Banking sector, there also is one organization who is largely responsible for the increase after March 1st. These two organizations are NOT included in the samples we use for our actual estimations. We do mention these observations in the main text as well.

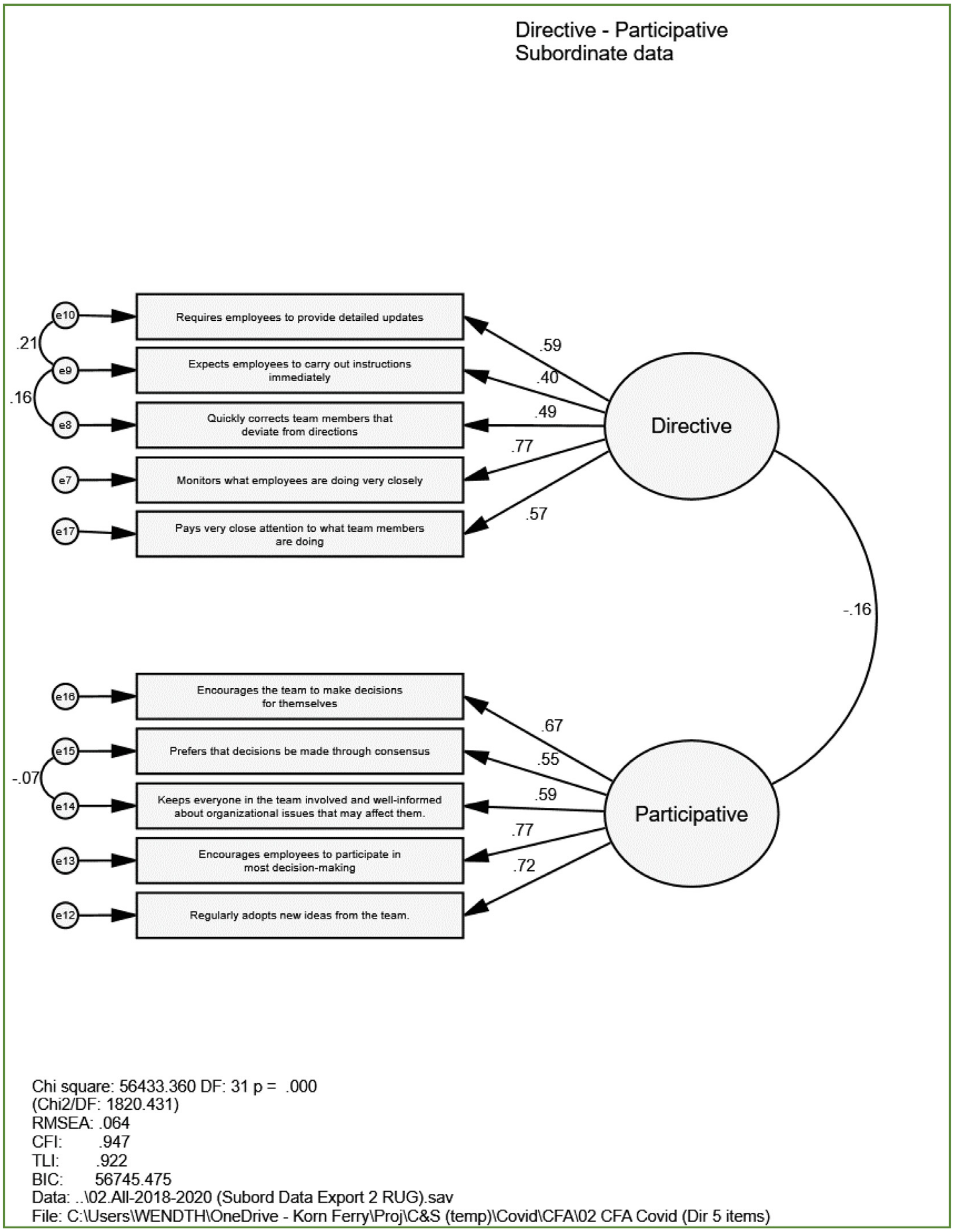
	Before March 1st Nr of Subordinates		**After March 1st** Nr of Subordinates			
	Sum	Column N %	Sum	Column N %		
Not coded	2,939		1,592		MAX%	Delta
Manufacturing	27,762	15.23%	4,694	6.93%	15.23%	−8.3%
Food Products	13,955	7.66%	3,097	4.57%	7.66%	−3.1%
Consumer Products (excluding Food & Beverage)	2,793	1.53%	1,218	1.80%	1.80%	0.3%
Chemical & Related Products	2,117	1.16%	594	0.88%	1.16%	−0.3%
Pharmaceuticals	12,497	6.86%	4,068	6.00%	6.86%	−0.9%
Technology	20,527	11.26%	10,476	15.46%	15.46%	4.2%
Telecommunications	1,361	0.75%	347	0.51%	0.75%	−0.2%
Financial Services	5,727	3.14%	4,627	6.83%	6.83%	3.7%
Banks/S&L's	26,990	14.81%	20,663	30.49%	30.49%	15.7%
Insurance	10,443	5.73%	2,500	3.69%	5.73%	−2.0%
Health	3,557	1.95%	239	0.35%	1.95%	−1.6%
Utilities	1,812	0.99%	419	0.62%	0.99%	−0.4%
Construction	3,573	1.96%	1,901	2.81%	2.81%	0.8%
Diversified Conglomerates	2,356	1.29%	2,875	4.24%	4.24%	3.0%
Agriculture	812	0.45%	145	0.21%	0.45%	−0.2%
Petroleum	4,419	2.42%	1,265	1.87%	2.42%	−0.6%
Mining	796	0.44%	97	0.14%	0.44%	−0.3%
Real Estate	4,670	2.56%	1,106	1.63%	2.56%	−0.9%
Retail	2,507	1.38%	1,012	1.49%	1.49%	0.1%
Hospitality and Tourism	97	0.05%			0.05%	−0.1%
Entertainment/Recreation	416	0.23%	127	0.19%	0.23%	0.0%
Wholesale Trade	177	0.10%	24	0.04%	0.10%	−0.1%
Transportation	1,696	0.93%	557	0.82%	0.93%	−0.1%
Communications	365	0.20%	330	0.49%	0.49%	0.3%
Broadcast Media	1,066	0.58%			0.58%	−0.6%
Professional Services	11,812	6.48%	2,035	3.00%	6.48%	−3.5%
Legal	609	0.33%	67	0.10%	0.33%	−0.2%
Professional Services – 3rd Parties	3,849	2.11%	473	0.70%	2.11%	−1.4%
Education	2,115	1.16%	714	1.05%	1.16%	−0.1%
Public Administration	6,157	3.38%	912	1.35%	3.38%	−2.0%
State & Local	3,603	1.98%	865	1.28%	1.98%	−0.7%
Associations	307	0.17%	151	0.22%	0.22%	0.1%
Miscellaneous	1,335	0.73%	165	0.24%	0.73%	−0.5%

Appendix D, sub b: Confirmatory Factor Analysis (CFA)

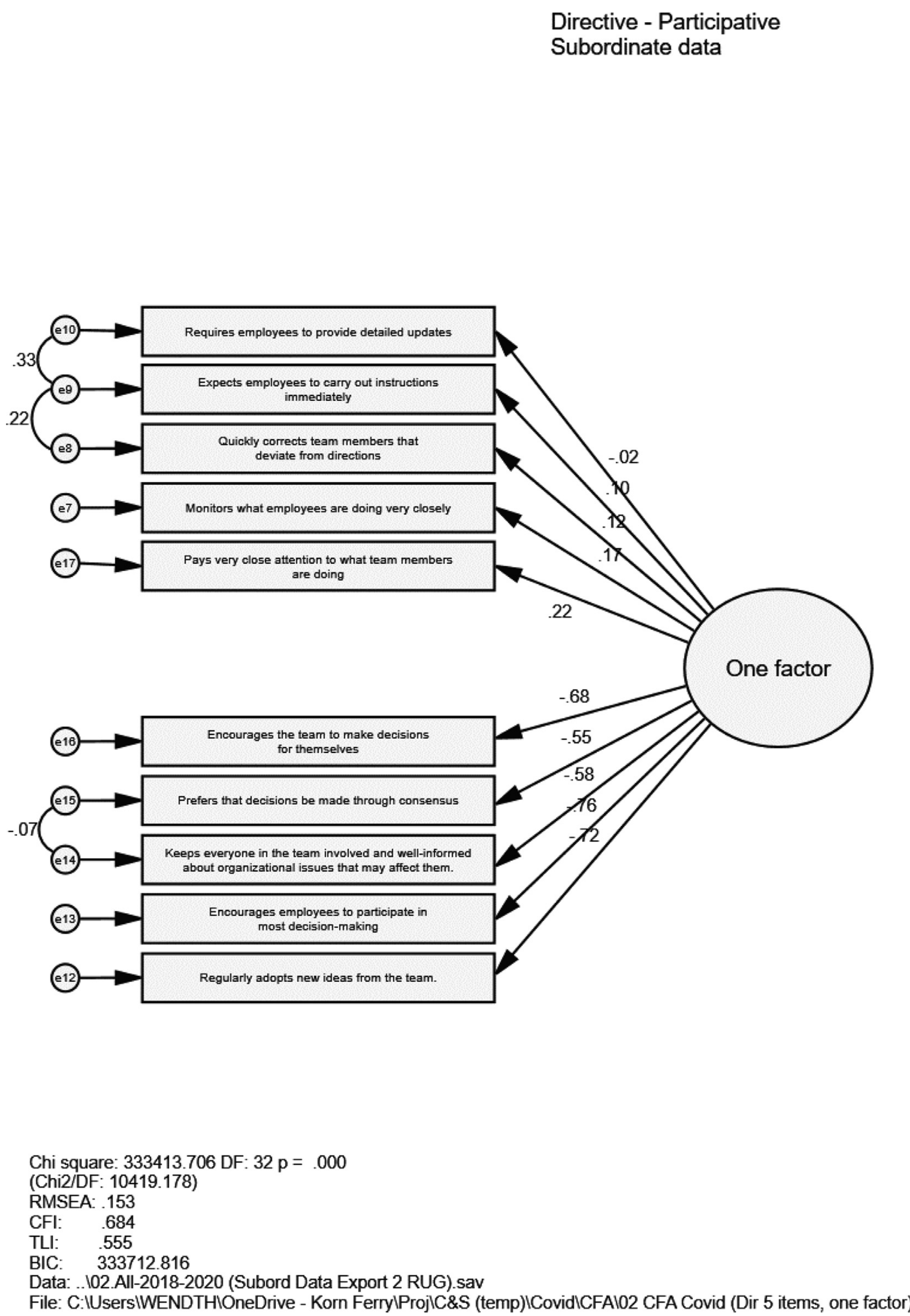
We ran a CFA analysis for a one and two-factor model, based on the original scales of KF. In these original scales, there were six items for directive leadership, and five for participative leadership. The CFA using all six items for directive (and five for participative) did not have a good fit: RMSEA > 0.1 and CFA, TLI < 0.9. Results were a little bit better by aggregating scores to the manager-ID, but still not satisfactory. Results are presented below:



In the next step, we therefore dropped the item “Makes most decisions for the people in the team” from the directive leadership scale. The result now was satisfactory: RMSEA 0.06 and CFI, TLI are all acceptable (see below). Moreover, it made conceptual sense to drop this item from our scale and arriving at this five-item scale is in line with another publication about the KF-dataset that uses directive leadership, namely Riisla et al. (2021).



Clearly, the two factor solution is better compared to the one factor solution (see below), according to the BIC: 56,745 versus 333,712 (https://www.methodology.psu.edu/resources/AIC-vs-BIC/). BIC is an estimate of a function of the posterior probability of a model being true, under a certain Bayesian setup, so that a lower BIC means that a model is considered to be more likely to be the true model.



Appendix D, sub c: Main Estimation Results for Rwg > 0.7 samples, Table 3A.

	Directive leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.133***	0.138***	0.112***	0.036	0.204***	0.163***	0.095*
	(0.032)	(0.032)	(0.039)	(0.064)	(0.061)	(0.042)	(0.052)
Management low	0.033*	0.020	0.039*	0.031	0.039*	0.039*	0.039*
	(0.019)	(0.019)	(0.021)	(0.022)	(0.021)	(0.021)	(0.021)
Management mid	−0.007	−0.017	−0.020	−0.023	−0.021	−0.021	−0.020
	(0.018)	(0.018)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)
Female	0.031*	0.035**	0.037**	0.037**	0.037**	0.036**	0.037**
	(0.016)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Age	−0.038***	−0.042***	−0.032***	−0.032***	−0.031***	−0.032***	−0.032***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Native	−0.062***	−0.066***	−0.123***	−0.124***	−0.124***	−0.122***	−0.123***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
WFHP high		−0.092***	−0.067***	−0.066***	−0.057***	−0.064***	−0.067***
		(0.017)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
WFHP mid		0.005	−0.041	−0.042	−0.039	−0.040	−0.041
		(0.024)	(0.027)	(0.027)	(0.028)	(0.027)	(0.027)
GDP			−0.174	−0.176	−0.172	−0.175	−0.175
			(0.298)	(0.298)	(0.298)	(0.298)	(0.298)
Power distance			0.008**	0.008**	0.008**	0.008**	0.008**
			(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
deaths_pop			−0.004	−0.004	−0.004	−0.004	−0.004
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
March 1st: Management low				0.175*			
				(0.092)			
March 1st:Management mid				0.051			
				(0.097)			
March 1st:WFHP high					−0.184**		
					(0.084)		
March 1st:WFHP mid					−0.053		
					(0.127)		
March 1st:deaths_pop						0.004***	
						(0.001)	
March 1st:power distance							−0.002
							(0.003)
Constant	3.539***	3.595***	5.152	5.178	5.130	5.158	5.161
	(0.085)	(0.086)	(3.162)	(3.159)	(3.162)	(3.155)	(3.161)
Observations	8,748	8,748	6,709	6,709	6,709	6,709	6,709
R^2							
Adjusted R^2							
Log Likelihood	−9,046.210	−9,036.694	−6,688.333	−6,689.433	−6,688.609	−6,688.937	−6,693.044
Akaike Inf. Crit.	18,110.420	18,095.390	13,404.670	13,410.870	13,409.220	13,407.880	13,416.090
Bayesian Inf. Crit.	18,174.110	18,173.230	13,500.020	13,519.840	13,518.200	13,510.040	13,518.260

Note: *p**p***p < 0.01.

Appendix D, sub c: Main Estimation Results for Rwg > 0.7 samples, Table 3B.

	Participative leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	−0.012	−0.016	−0.017	−0.009	−0.054	−0.035	−0.017
	(0.033)	(0.033)	(0.033)	(0.054)	(0.051)	(0.035)	(0.043)
Management low	−0.008	0.003	0.002	0.006	0.002	0.002	0.002
	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Management mid	−0.017	−0.009	−0.009	−0.011	−0.008	−0.008	−0.009
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Female	0.083***	0.080***	0.080***	0.080***	0.081***	0.081***	0.080***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age	0.008	0.011	0.010	0.011	0.010	0.010	0.010
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Native	0.058***	0.059***	0.061***	0.061***	0.061***	0.060***	0.061***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
WFHP high		0.056***	0.057***	0.057***	0.054***	0.056***	0.057***
		(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)
WFHP mid		0.015	0.011	0.013	0.008	0.011	0.011
		(0.022)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)
GDP			−0.034	−0.033	−0.035	−0.034	−0.034
			(0.103)	(0.103)	(0.103)	(0.102)	(0.103)
Power distance			−0.002	−0.002	−0.002	−0.002	−0.002
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
deaths_pop			0.001*	0.001*	0.001*	0.001**	0.001*
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
March 1st:Management low				−0.072			
				(0.077)			
March 1st:Management mid				0.061			
				(0.081)			
March 1st:WFHP high					0.059		
					(0.070)		
March 1st:WFHP mid					0.081		
					(0.105)		
March 1st:deaths_pop						−0.001	
						(0.001)	
March 1st:power distance							0.00000
							(0.003)
Constant	4.390***	4.350***	4.751***	4.735***	4.762***	4.744***	4.751***
	(0.047)	(0.048)	(1.087)	(1.087)	(1.089)	(1.085)	(1.087)
Observations	6,709	6,709	6,709	6,709	6,709	6,709	6,709
R^2							
Adjusted R^2							
Log Likelihood	−5,440.395	−5,439.708	−5,449.977	−5,451.993	−5,452.632	−5,455.042	−5,454.991
Akaike Inf. Crit.	10,898.790	10,901.420	10,927.950	10,935.990	10,937.260	10,940.080	10,939.980
Bayesian Inf. Crit.	10,960.090	10,976.340	11,023.310	11,044.970	11,046.240	11,042.250	11,042.150

Note: *p**p***p < 0.01.

Appendix D, sub c: Main Estimation Results for Rwg > 0.7 samples, Table 4A.

	Directive Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.025	0.023	0.014	0.004	0.012	0.013	0.011
	(0.016)	(0.016)	(0.016)	(0.025)	(0.021)	(0.016)	(0.016)
Management low	0.040**	0.031*	0.031*	0.026	0.031*	0.031*	0.032*
	(0.017)	(0.017)	(0.017)	(0.019)	(0.017)	(0.017)	(0.017)
Management mid	−0.001	−0.008	−0.009	−0.013	−0.009	−0.008	−0.008
	(0.016)	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)	(0.016)
Female	0.020	0.023	0.023	0.023	0.023	0.023	0.023
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Age	−0.031***	−0.034***	−0.033***	−0.033***	−0.033***	−0.033***	−0.033***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)

(continued on next page)

Appendix D (*continued*)

	Directive Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Native	−0.040*	−0.045**	−0.045**	−0.045**	−0.045**	−0.045**	−0.045**
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
WFHP high		−0.076***	−0.077***	−0.077***	−0.079***	−0.077***	−0.077***
		(0.016)	(0.016)	(0.016)	(0.018)	(0.016)	(0.016)
WFHP mid		0.008	0.008	0.007	0.009	0.007	0.007
		(0.022)	(0.022)	(0.022)	(0.024)	(0.022)	(0.022)
GDP			−0.172***	−0.171***	−0.172***	−0.170***	−0.174***
			(0.065)	(0.065)	(0.065)	(0.065)	(0.065)
Power distance			0.009***	0.009***	0.009***	0.009***	0.009***
			(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Deaths_pop			−0.002**	−0.002**	−0.002**	−0.002**	−0.002**
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
March 1st:Management low				0.019			
				(0.039)			
March 1st:Management mid				0.016			
				(0.036)			
March 1st:WFHP high					0.009		
					(0.035)		
March 1st:WFHP mid					−0.008		
					(0.050)		
March 1st:Deaths_pop						0.00004	
						(0.0002)	
March 1st:Power distance							−0.001
							(0.001)
Constant	3.485***	3.530***	5.219***	5.219***	5.225***	5.207***	5.240***
	(0.081)	(0.082)	(0.647)	(0.647)	(0.648)	(0.650)	(0.649)
Observations	10,888	10,888	10,888	10,888	10,888	10,888	10,888
Log Likelihood	−11,250.090	−11,242.960	−11,235.370	−11,240.060	−11,239.850	−11,243.070	−11,240.600
Akaike Inf. Crit.	22,518.170	22,507.930	22,498.740	22,512.120	22,511.700	22,516.140	22,511.190
Bayesian Inf. Crit.	22,583.830	22,588.170	22,600.870	22,628.850	22,628.430	22,625.570	22,620.630

Appendix D, sub c: Main Estimation Results for Rwg > 0.7 samples, Table 4B.

	Participative Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
March 1st	0.059***	0.061***	0.056***	0.059***	0.042***	0.056***	0.056***
	(0.009)	(0.009)	(0.010)	(0.015)	(0.012)	(0.010)	(0.010)
Management low	−0.007	0.003	0.004	0.004	0.004	0.004	0.004
	(0.010)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Management mid	−0.022**	−0.016	−0.016	−0.013	−0.016	−0.016	−0.016
	(0.010)	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)
Female	0.043***	0.040***	0.041***	0.041***	0.041***	0.041***	0.040***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age	0.012**	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
native	0.020	0.020	0.019	0.019	0.020	0.019	0.019

Appendix D (*continued*)

	Participative Leadership						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
WFHP high		0.062***	0.063***	0.063***	0.053***	0.063***	0.063***
		(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
WFHP mid		0.042***	0.041***	0.041***	0.033**	0.041***	0.041***
		(0.013)	(0.013)	(0.013)	(0.015)	(0.013)	(0.013)
GDP			−0.110***	−0.110***	−0.110***	−0.110***	−0.111***
			(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Power distance			−0.001	−0.001	−0.001	−0.001	−0.001
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Deaths_pop			0.001***	0.001***	0.001***	0.001***	0.001***
			(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
March 1st:Management low				0.003			
				(0.023)			
Marcch 1st:Management mid				−0.011			
				(0.022)			
March 1st:WFHP high					0.041*		
					(0.022)		
March 1st:WFHP mid					0.030		
					(0.031)		
March 1st:Deaths_pop						0.00001	
						(0.0001)	
March 1st:Power distance							−0.001
							(0.001)
Constant	4.684***	4.649***	5.741***	5.741***	5.744***	5.740***	5.743***
	(0.032)	(0.033)	(0.232)	(0.232)	(0.232)	(0.232)	(0.232)
Observations	15,799	15,799	15,799	15,799	15,799	15,799	15,799
Log Likelihood	−11,716.190	−11,702.700	−11,704.750	−11,710.410	−11,708.310	−11,712.970	−11,710.850
Akaike Inf. Crit.	23,450.370	23,427.400	23,437.500	23,452.820	23,448.630	23,455.930	23,451.710
Bayesian Inf. Crit.	23,519.380	23,511.750	23,544.850	23,575.500	23,571.310	23,570.940	23,566.720

Appendix D sub d: Additional sample comparisons

(I) Total sample before/after March 1st 2020;
(II) for WFH and level of management sub-samples.

(I) Total sample.
The sample sizes are different: the sub-sample of the crisis period (= after March 1st 2020) is much smaller than de pre-crisis sample, especially because of the timeframe. The results below show that the two sub-samples are however very comparable on the individual.

variables (management, gender, age and nativeness); note that these scores are thus also very much in line with the respective descriptive statistics in Table 1 (for the total sample).

Sample	Before March 1st	After March 1st
Observations	21,651	6902
Mean Management (and SD)	2 (0.82)	2.1 (0.81)
Mean Gender (and SD)	0.3 (0.46)	0.25 (0.43)
Mean Age (and SD)	3.9 (0.84)	3.9 (0.81)
Mean Native (and SD)	0.88 (0.33)	0.9 (0.3)

(II) Sub-sample comparisons for WFHP and level of management categories.

See below the results of the sample before versus after March 1st, when we look at sub-samples of the three WFHP categories and the three management categories. Again, they are very comparable.

Variable	**WFHP Low**		**WFHP High**		**WFHP Mid**	
	Before March 1st	*After March 1st*	*Before March 1st*	*After March 1st*	*Before March 1st*	*After March 1st*
Age	3.90	3.80	3.90	3.98	4.04	4.13
Native	0.89	0.92	0.83	0.83	0.94	0.92
Gender	0.26	0.24	0.37	0.28	0.32	0.28
Management	2.05	1.93	1.93	1.91	1.90	1.86
Variable	**Low Management**		**Mid Management**		**High Management**	
	Before March 1st	*After March 1st*	*Before March 1st*	*After March 1st*	*Before March 1st*	*After March 1st*
Age	3.68	3.57	3.94	3.92	4.14	4.04
Native	0.87	0.90	0.88	0.92	0.88	0.89
Gender	0.36	0.31	0.28	0.22	0.26	0.24
WFHP	1.49	1.33	1.55	1.46	1.68	1.49

Appendix D sub e: Correlation table in levels

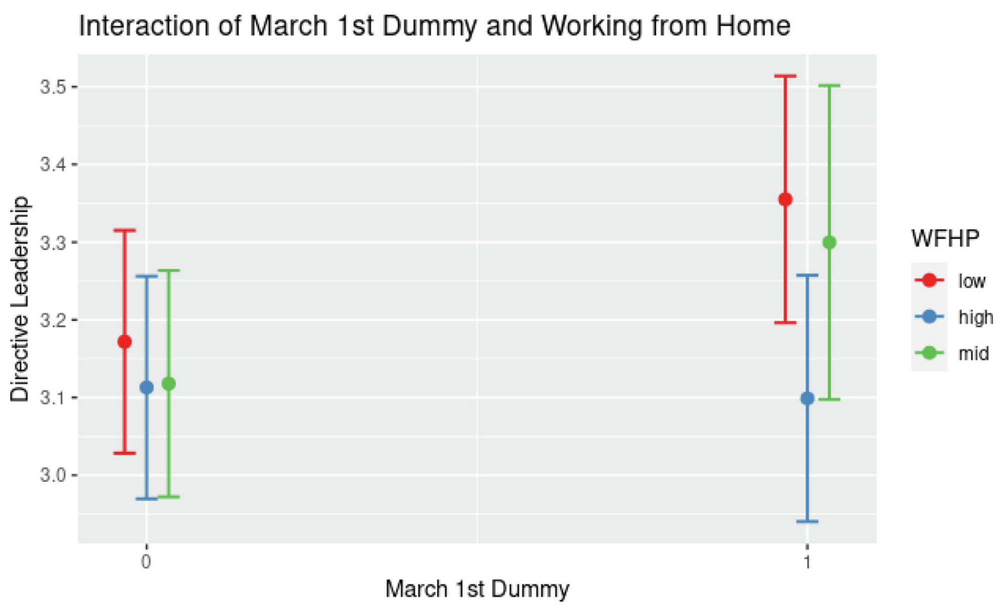
	Gender	Age	Native	WFHP low	WFHP high	WFHP mid	Mgt high	Mgt low	Mgt mid
Gender	1								
Age	−0.08***	1							
Native	−0.02***	−0.02***	1						
WFHP low	−0.09***	−0.05***	0.05***	1					
WFHP high	0.08***	0	−0.1***	−0.74***	1				
WFHP mid	0.02***	0.07***	0.06***	−0.47***	−0.24***	1			
Management high	−0.06***	0.18***	0	−0.09***	0.05***	0.06***	1		
Management low	0.1***	−0.21***	−0.02***	0.08***	−0.04***	−0.06***	−0.51***	1	
Management mid	−0.04***	0.02***	0.01***	0.01*	−0.01	0	−0.51***	−0.48***	1

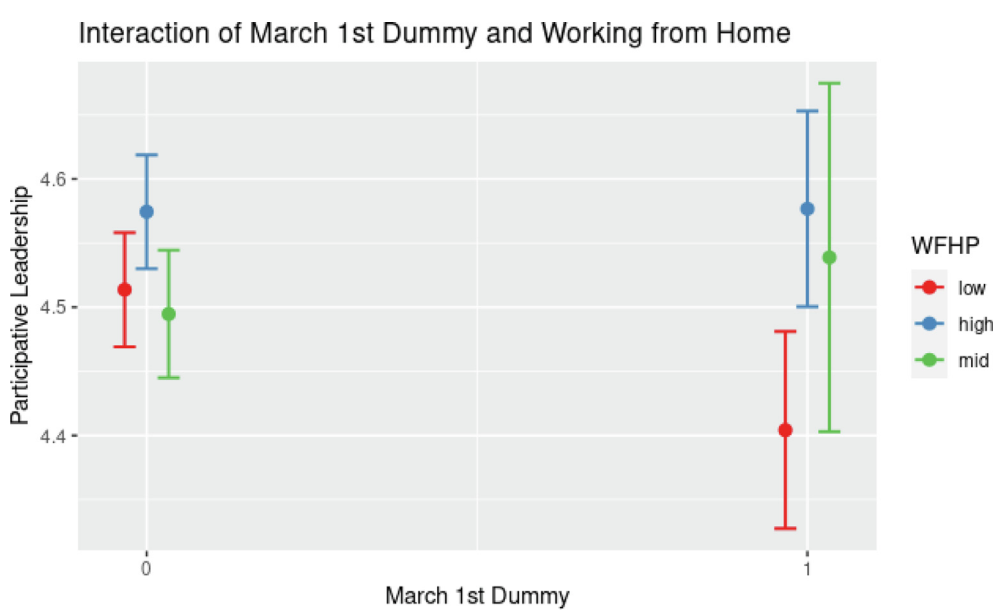
N = 28554

	GDP	Power distance	Deaths (per 100 K)
GDP	1		
Power Distance	−0.73***	1	
Deaths (per 100 K)	0.06	−0.08	1

N = 48

Appendix D sub f: Confidence intervals for bar plots for interaction of the March 1st dummy and WFHP, for directive leadership and participative leadership





References

Aguinis, H., Boyd, B. K., Pierce, C. A., & Short, J. C. (2011). Walking new avenues in management research methods and theories: Bridging micro and macro domains. *Journal of Management, 37*, 395–403.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics*. Princeton University Press.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2014). Causality and endogeneity: Problems and solutions. *The Oxford handbook of leadership and organizations, 1*, 93–117.

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5–21.

Antonakis, J., Banks, G., Bastardoz, N., Cole, M., Day, D., Eagly, A., ... Weber, R. (2019). The leadership quarterly: State of the journal. *The Leadership Quarterly, 30*(1), 1–9.

Antonakis, J., Bastardoz, N., & Rönkkö, M. (2021). On ignoring the Random Effects Assumption in Multilevel Models: Review, Critique, and Recommendations. *Organizational Research Methods, 24*(2), 443–483.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*, 1086–1120.

Beugelsdijk, S., Maseland, R., & Van Hoorn, A. (2015). Are Scores on H ofstede's Dimensions of National Culture Stable over Time? *A Cohort Analysis. Global Strategy Journal, 5*(3), 223–240.

Bliese, P. D., Maltarich, M. A., & Hendricks, J. L. (2018). Back to basics with mixed-effects models: nine take-away points. *Journal of Business and Psychology, 33*(1), 1–23.

Bonet, R., & Salvador, F. (2017). When the boss is away: Manager–worker separation and worker performance in a multisite software maintenance organization. *Organization Science, 28*(2), 244–261.

Card, D. (1992). Using regional variation in wages to measure the effects of the federal minimum wage. *Ilr Review, 46*(1), 22–37.

Chattopadhyay, P., Glick, W. H., & Huber, G. P. (2001). Organizational actions in response to threats and opportunities. *Academy of Management Journal, 44*, 937–955.

Chun, J. U., Yammarino, F. J., Dionne, S. D., Sosik, J. J., & Moon, H. K. (2009). Leadership across hierarchical levels: Multiple levels of management and multiple levels of analysis. *The Leadership Quarterly, 20*(5), 689–707.

Clair St., T., & Cook, T. D. (2015). Difference-in-differences methods in public finance. *National Tax Journal, 68*(2), 319–338.

Contreras, F., Baykal, E., & Abid, G. (2020). E-leadership and teleworking in times of COVID-19 and beyond: What we know and where do we go. *Frontiers in Psychology, 11*, 3484.

Dickson, M. W., Den Hartog, D. N., & Mitchelson, J. K. (2003). Research on leadership in a cross-cultural context: Making progress, and raising new questions. *The Leadership Quarterly, 14*, 729–768.

Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home? Journal of Public Economics, Vol. 189, forthcoming, see also https://www.nber.org/papers/w26948.

Eckardt, R., Crocker, A., & Tsai, C. Y. (2020). Clarifying and empirically assessing the concept of human capital resource emergence. *International Journal of Human Resource Management, 32*(2), 279–306.

Euwema, M. C., Wendt, H., & Van Emmerik, H. (2007). Leadership styles and group organizational citizenship behavior across cultures. *Journal of Organizational Behavior, 28*, 1035–1057.

Franchini, A. F., Auxilia, F., Galimberti, P. M., Piga, M. A., Castaldi, S., & Porro, A. (2020). COVID 19 and Spanish flu pandemics: All it changes, nothing changes. *Acta Bio Medica: Atenei Parmensis, 91*(2), 245.

Garretsen, H., Stoker, J. I., & Weber, R. A. (2020). Economic perspectives on leadership: Concepts, causality, and context in leadership research. *The Leadership Quarterly, 31* (3) 101410.

Gilbert, C. G. (2005). Unbundling the structure of inertia: Resource versus routine rigidity. *Academy of Management Journal, 48*, 741–763.

Gladstein, D. L., & Reilly, N. P. (1985). Group decision making under threat: The tycoon game. *Academy of Management Journal, 28*, 613–627.

Hale, T., Petherick, A., Phillips, T., & Webster, S. (2020). Variation in government responses to COVID-19. *Blavatnik school of government working paper*, *31*, 2020-11.

Hiller, N. J., Piccolo, R. F., & Zaccaro, S. J. (2020). Economic assumptions and economic context: Implications for the study of leadership. *The Leadership Quarterly, 31*(3) 101352.

Hlavac, M. (2015). ***Stargazer: Well-formatted regression and summary statistics tables***. R package version5, 2. URL: http://CRAN.R-project.org/package=stargazer.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills CA: Sage Publications.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage.

Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research, 38*, 52–88.

House, R. J. (1971). A path goal theory of leader effectiveness. *Administrative Science Quarterly, 16*, 321–339.

House, R., Javidan, M., Hanges, P., & Dorfman, P. (2002). Understanding cultures and implicit leadership theories across the globe: An introduction to project GLOBE. *Journal of World Business, 37*, 3–10.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.

IMF (2020). *World Economic Outlook, October 2020*. Washington DC: International Monetary Fund.

Jacquart, P., & Antonakis, J. (2015). When does charisma matter for top-level leaders? Effect of attributional ambiguity. *Academy of Management Journal, 58*, 1051–1074.

Jacquart, P., Santoni, S., Schudy, S., Sieweke, J., & Withers, M. C. (2020). Special Issue on Harnessing Exogenous Shocks for Leadership and Management Research. *The Leadership Quarterly*, *31*(5), 101464.

Jackson, T. A., Meyer, J. P., & Wang, X. H. F. (2013). Leadership, commitment, and culture: A meta-analysis. *Journal of Leadership and Organizational Studies, 20*, 84–106.

Jago, A. G., & Vroom, V. H. (1977). Hierarchical level and leadership style. *Organizational Behavior and Human Performance, 18*(1), 131–145.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*, 219–229.

Johns, G. (2018). Advances in the treatment of context in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior, 5*, 21–46.

Kamphuis, W., Gaillard, A. W., & Vogelaar, A. L. (2011). The effects of physical threat on team processes during complex task performance. *Small Group Research, 42*, 700.

Kirchner, K., Ipsen, C., & Hansen, J. P. (2021). COVID-19 leadership challenges in knowledge work. *Knowledge Management Research & Practice, 1–8*, 729.

Kirkman, B. L., Chen, G., Farh, J. L., Chen, Z. X., & Lowe, K. B. (2009). Individual power distance orientation and follower reactions to transformational leaders: A cross-level, cross-cultural examination. *Academy of Management Journal, 52*(4), 744–764.

Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. J., Bakker, A. B., ... Creary, S. J. (2021). COVID-19 and the Workplace: Implications, Issues, and Insights for Future Research and Action. *American Psychologist, 76*(1), 63–77.

Lorinkova, N. M., Pearsall, M. J., & Sims, H. P. (2013). Examining the differential longitudinal performance of directive versus empowering leadership in teams. *Academy of Management Journal, 56*, 573–596.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *The Leadership Quarterly, 7*, 385–425.

Madera, J. M., & Smith, D. B. (2009). The effects of leader negative emotions on evaluations of leadership in a crisis situation: The role of anger and sadness. *The Leadership Quarterly, 20*(2), 103–114.

Meyer, A. D. (1982). Adapting to environmental jolts. *Administrative Science Quarterly, 27*(4), 515–537.

Moorty, L., Mahajan, S., Andronova, V. G., & Izvorski, I. (2020). *A Policy Framework for Mitigating the Impact of the COVID-19 Crisis*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/33776.

Morens, D. M., Daszak, P., Markel, H., & Taubenberger, J. K. (2020). Pandemic COVID-19 joins history's pandemic legion. *Mbio, 11*(3).

Oc, B. (2018). Contextual leadership: A systematic review of how contextual factors shape leadership and its outcomes. *The Leadership Quarterly, 29*(1), 218–235.

Oshagbemi, T. (2008). The impact of personal and organisational variables on the leadership styles of managers. *The International Journal of Human Resource Management, 19*, 1896–1910.

Oshagbemi, T., & Gill, R. (2004). Differences in leadership styles and behaviour across hierarchical levels in UK organisations. *Leadership & Organization Development Journal, 25*, 93–106.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Core Team, R. (2015). ***Nlme: Linear and nonlinear mixed effects models.*** R package version3, 1–122. URL: http://CRAN.R-project.org/package=nlme.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. In J. B. Taylor & H. Uhlig (Eds.). *Handbook of Macroeconomics* (Vol. 2, pp. 71–162). Elsevier.

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Riisla, K., Wendt, H., Babalola, M. T., & Euwema, M. (2021). Building Cohesive Teams—The Role of Leaders' Bottom-Line Mentality and Behavior. *Sustainability, 13*(14), 8047.

Rudolph, C. W., Allan, B., Clark, M., Hertel, G., Hirschi, A., Kunze, F., ... Zacher, H. (2021). Pandemics: Implications for research and practice in industrial and organizational psychology. *Industrial and Organizational Psychology, 14*(1–2), 1–35.

Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science Advances, 7*(21), eabd1705.

Shangguan, Z., Wang, M. Y., & Sun, W. (2020). What caused the outbreak of COVID-19 in China: From the perspective of crisis management. *International Journal of Environmental Research and Public Health, 17*(9), 3279.

Shimizu, K. (2007). Prospect theory, behavioral theory, and the threat-rigidity thesis: Combinative effects on organizational decisions to divest formerly acquired units. *Academy of Management Journal, 50*, 1495–1514.

Sieweke, J., & Santoni, S. (2020). Natural experiments in leadership research: An introduction, review, and guidelines. *The Leadership Quarterly, 31*(1) 101338.

Somech, A. (2005). Directive versus participative leadership: Two complementary approaches to managing school effectiveness. *Educational Administration Quarterly, 41*, 777–800.

Somech, A. (2006). The effects of leadership style and team process on performance and innovation in functionally heterogeneous teams. *Journal of Management, 32*, 132–157.

Staw, B. M., Sandelands, L. E., & Dutton, J. E. (1981). Threat-rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly, 26*, 501–524.

Stoker, J. I., Garretsen, H., & Lammers, J. (2022). Leading and working from home in times of COVID-19: On the perceived changes in leadership behaviors. *Journal of Leadership and Organizational Studies, 29*(2), 208–218.

Stoker, J. I., Garretsen, H., & Soudis, D. (2019). Tightening the leash after a threat: A multi-level event study on leadership behavior following the financial crisis. *The Leadership Quarterly, 30*, 199–214.

Tuncdogan, A., Acar, O. A., & Stam, D. (2017). Individual differences as antecedents of leader behavior: Towards an understanding of multi-level outcomes. *The Leadership Quarterly, 28*, 40–64.

Van Emmerik, H., Wendt, H., & Euwema, M. C. (2010). Gender ratio, societal culture, and male and female leadership. *Journal of Occupational and Organizational Psychology, 83*, 895–914.

Van Mierlo, H., Vermunt, J., & Rutte, C. C. (2009). Composjng group-level constructs from individual-level survey data. *Organizational Research Methods, 12*(2), 368–392.

Yukl, G. A. (2013). *Leadership in Organizations* (8th ed.). Albany: State University of New York.