UNIVERSITÄT PADERBORN

HEINZ NIXDORF INSTITUT
UNIVERSITÄT PADERBORN

Perspectives on Social Norms and Social Responsibility:

Emerging Economic and Managerial Approaches

KUMULATIVE DISSERTATION

Fakultät für Wirtschaftswissenschaften der

Universität Paderborn.

zur Erlangung des akademischen Grades „Doktor der Wirtschaftswissenschaften“

Doctor rerum politicarum (Dr. rer. pol.)

vorgelegt von

Sabrina Plaß, M.Sc.

Geboren am 27. Juli 1990 in Lingen (Ems)

Eingereicht im April 2025

Danksagung

Diese Arbeit wäre ohne meine Wegbegleiter:innen nicht möglich gewesen. An dieser Stelle möchte ich mich deshalb bei all jenen bedanken, die mich in den vergangenen Jahren auf unterschiedliche Weise – fachlich und/oder emotional – unterstützt haben.

Zunächst ein herzliches Dankeschön an meine Familie, Hannelore und Reinhard, Anna und Marcel mit Chiara und Katharina, die stets Vertrauen und Verständnis hatten. Ebenso an meine Freund:innen für das gemeinsame Erleben der Höhen und Tiefen dieser Zeit. Dieser kurze Absatz wird euch nicht gerecht, doch alles Weitere würde hier den Rahmen sprengen.

Danke an alle, die mich schon vor Beginn der Promotion darin bestärkt haben, diesen Weg zu gehen – Dörte, Mark, David, um nur einige zu nennen. Ein besonderer Dank gilt auch Sven, der mir in den letzten Monaten mit seiner fröhlichen Art viel Energie für den Endspurt gegeben hat.

Ein großes Dankeschön an meine Ko-Autor:innen und Kolleg:innen: Danke Behnud und Sabrina für die gemeinsamen Paper-Entwicklungs-Achterbahnfahrten. Danke an Eva, mit der ich das erste Paper in nun dutzenden Schleifen weiterentwickeln durfte. Danke an das gesamte Dual-Strat-Team für die gemeinsame Arbeit an den Projekten. Vielen Dank an Marius, der stets ein offenes Ohr für meine Statistik- und Experiment-Fragen hatte. Lisa, Dörte und Tanya – danke für die tolle Zusammenarbeit bei der Unterstützung der Lehre im Kontext Nachhaltigkeit. Ein herzliches Dankeschön an alle Kolleg:innen, die mir in den letzten Monaten den Rücken freigehalten haben, für die fachliche Unterstützung, die schönen Gespräche, die gemeinsamen Mittagspausen und den Spaß bei und neben der Arbeit, der wesentlich dazu beigetragen hat, immer wieder Motivation zu finden. Gleichermaßen ein großes Danke an Melanie für ihre herzliche Art und die Unterstützung bei all den bürokratischen Angelegenheiten! Und natürlich ein herzliches Dankeschön an René, der mir überhaupt erst die Möglichkeit und das Vertrauen sowie die Offenheit für Gestaltung gegeben hat und mir damit all diese wunderbaren und lehrreichen Erfahrungen am Lehrstuhl ermöglichte. Mein Dank gilt außerdem meiner Prüfungskommission für die Begleitung dieses Promotionsvorhabens.

Table of Contents

CHAPTER 1 | Synopsis

Ecological crises, social unrest, and digital transformation have raised the need to rethink the economy. As such, digitalisation is transforming business operations and redefining their responsibilities towards society and the environment. Corporate self-commitment has been debated and practised for decades, yet its effectiveness depends on moving beyond polished mission statements and embedding sustainability into core business operations. Fuelled by external drivers – such as regulatory pressures, societal expectations, and technological advances – sustainability, once considered an add-on or ethical obligation, is increasingly becoming a strategic requirement for corporations. The pervasiveness of digital technologies requires managers to make decisions that balance economic, environmental, and social impacts of digitalisation, while leveraging digital technologies for genuine sustainability.

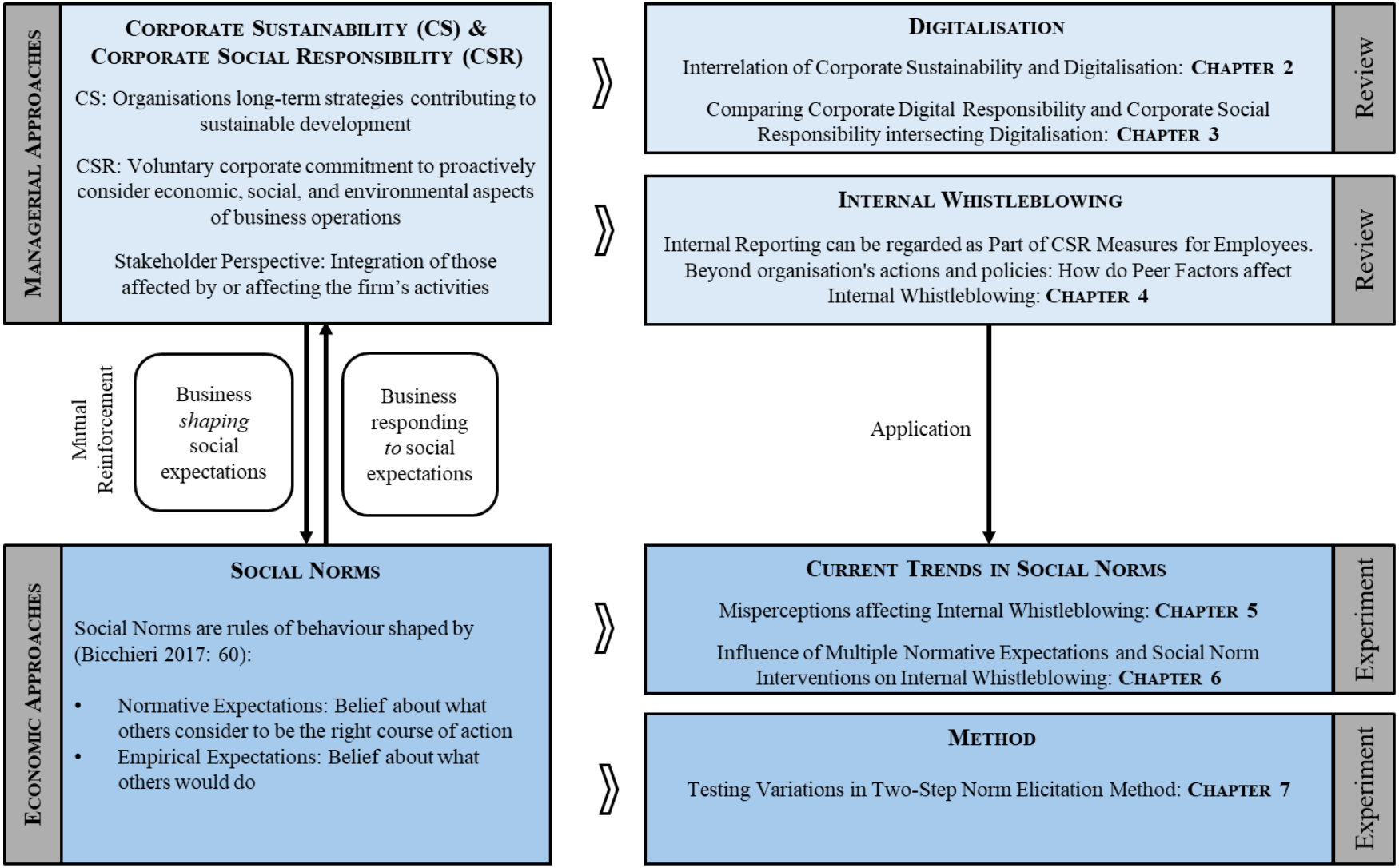
Amid the megatrends of sustainability and digitalisation, scandals like Facebook-Cambridge Analytica or Wirecard expose the dark side of corporate social irresponsibility. This emphasises a long-standing question: How should individuals behave when they observe corporate misconduct? CSR and whistleblowing are profoundly interrelated: While whistleblowing exposes corporate social irresponsibility, implementing whistleblowing mechanisms within credible CSR initiatives signals a genuine commitment to CSR practices and ensures that companies 'walk their talk'. However, successful organisational institutionalisation is only one side of the coin and does not guarantee whistleblowers will come forward. In this context, the influence of colleagues and social norms is increasingly apparent.

The first set of articles in this dissertation critically examines the interplay between Corporate Sustainability (CS) and digitalisation, questioning how digital advancements align with sustainable corporate practices. It then introduces the emerging concept of Corporate Digital Responsibility (CDR), challenging its place alongside the more established Corporate Social Responsibility (CSR). Finally, it examines how coworkers influence the likelihood of potential whistleblowers coming forward. Building on these considerations, the second set of articles employs behavioural economic experiments to explore the nuanced social expectations that influence whistleblowing on the micro level. These studies investigate how individuals (mis)perceive their peers'

expectations: Do potential whistleblowers correctly estimate whether others believe it is appropriate to speak up? How do conflicting social expectations of two mutually exclusive actions – reporting and staying silent – as well as norm interventions affect decision-making? And how robust are the methods used to measure these normative expectations?

This dissertation maps the evolving landscape of CS and CSR in the digital age and outlines the mechanisms that shape individuals' whistleblowing decisions. The findings aim to contribute to a deeper understanding of how organisations can be sustainable and digitally progressive and how the social environment shapes our ethical decision-making. Figure 1 on the next page presents a simplified overview of the chapters, which will be outlined in more detail in the next section.

Figure 1: Simplified Visual Illustration of Chapters



1.1 CS and CSR Evolving in the Context of Digitalisation

In the 21st century, we are part of transformations in (at least) two perspectives. On the one hand, we desperately need change to achieve sustainable development. This includes rethinking our ways of production and consumption to mitigate accumulated ecological and social crises. On the other hand, we witness tremendous change (and speed) in how digitalisation interferes with and disrupts economies, businesses and everyday life. Both megatrends, sustainability and digitalisation, are talked about a lot and sometimes referred to as wicked problems, which are characterised as multi-layered, challenging to define and resolve, and subject to vested interest and outcomes requiring diverse stakeholder collaborations (Elliott & Copilah-Ali, 2024a). The interaction of both can create synergies and rebound effects (Santarius, 2016). For instance, remote work may reduce the carbon footprint regarding staff transportation and office energy costs; however, it may increase the digital carbon footprint due to more reliance on cloud services, video conferences, and energy costs at home. The ambivalent effects of technology with regard to planetary boundaries have been debated since the Limits to Growth report (1972) presented by the Club of Rome. After all, scholars increasingly recognise the need to dovetail these megatrends in a dual transformation (Epp et al., 2024; Kürpick et al., 2024) or twin transformation (Barth et al., 2023). The following paragraph will more closely reflect on the sustainability perspective and then outline the concepts of corporate sustainability and corporate social responsibility.

From an ecological-economic perspective, sustainability entails a development that aligns economic activities with the use and absorption of natural resources in a socially responsible manner (Daly, 2007). Such a holistic view of sustainability regards economic interdependencies and refers to all economic activities as a subsystem of the social foundation and the superordinate ecological system. In this realm, natural resource substitutability is regarded as limited, with non-renewable natural resources and other production factors often being complements (DesJardins, 2007; Giddings et al., 2002). One alternative approach taking all these considerations into account is the so-called model of the doughnut economy. It points to economic prosperity being subject to the importance of a necessary social minimum, which is still marked by a shortfall of essential living standards by some, and ecological boundaries, which are increasingly overshot (Raworth, 2018).

From a managerial perspective, the challenges and potentials that are accompanied by sustainability and responsibility towards stakeholders are comprised, among others, in the following concepts which have proliferated during the past decades[1]: Corporate Sustainability (CS), describing a system perspective and long-term oriented corporate strategies; Corporate Social Responsibility (CSR), a business approach of companies being socially accountable to the society (Bansal & Song, 2017); and Stakeholder Theory (e.g., Freeman, 2010; Parmar et al., 2010), the alignment of business activities with those affected by or affecting the company. These concepts have different historical roots; however, they increasingly aim for a similar purpose (Bansal & Song, 2017; Dmytriyev et al., 2021; Montiel, 2008). While it is not the aim to trace back the development of these concepts – an endeavour already undertaken by scholars (e.g., Bansal & Song, 2017) – some commonalities and differences will be highlighted. As a commonality, all concepts describe a foremost voluntary approach of business to consider stakeholder and social expectations in corporate operations beyond legal requirements to contribute to the (sustainable) social good. The main differences are the communities and perspectives these concepts emerged from: CS developed since the 1980s and adopts a system perspective regarding corporations as an embedded system, predominated by environmental concerns, whereas CSR originates from normative discussions of good business conduct, predominated by ethical and social issues (Bansal & Song, 2017). In the last decades, these concepts converge more closely, integrating similar environmental, social and economic aspects (Bansal & Song, 2017; Montiel, 2008). Finally, Stakeholder Theory has been primarily a strategic management approach, suggesting that companies achieve long-term success by incorporating their stakeholders' needs and expectations and aiming to create value for their stakeholders (Freeman & Elms, 2023). The stakeholder perspective has become a common approach in CSR (McWilliams & Siegel, 2001). For instance, it incorporates expectations of and within the supply chain that are not purely efficiency-driven but build on long-term supplier cooperation. The stakeholder approach further manifests in the maturity assessment of sustainability reporting and strategy building, for which stakeholders' expectations are evaluated.

[1] Additional concepts have emerged as "offspring" concepts of CSR, for instance, Corporate Social Performance, Corporate Political Responsibility or Corporate Citizenship (Latapí Agudelo et al., 2019). Carroll (2021) critically reflects on the considerable overlap of these concepts with CSR.

If aligned with sustainability goals, digital technologies have, for instance, the potential to contribute to a circular economy. To more closely reflect on the current state of research on how CS and digitalisation are currently interrelated, Chapter 2 conducts a systematic literature review and derives corresponding thematic clusters. These clusters yield key insights: One is that efficiency-driven perspectives dominate the literature, focusing on the ecological and economic benefits of integrating sustainability and digitalisation; however, the emphasis on social and governance dimensions grows. Additionally, digitalisation is primarily framed as an enabler rather than an independent goal of sustainability, and strategic alignment is crucial for integrating digitalisation and sustainability within corporate practices (for an overview of studies in this dissertation, see Table 1, p. 14-15).

Digitisation is also increasingly being considered in the context of CSR. Debates centre about ethics and accountability in Artificial Intelligence (AI), big data, and algorithms (e.g., Coeckelbergh, 2020; Herschel & Miori, 2017; Mittelstadt et al., 2016). For instance, consumer scepticism over algorithmic decision-making and concerns about biases or data privacy emerged due to data security breaches (Bernstein, 2017; George et al., 2014; Newell & Marabelli, 2015). Therefore, digital trust is becoming an important asset for companies (Kluiters et al., 2023a). Additionally, virtual stakeholders have emerged (Freeman et al., 2017). For instance, influencers have become a new intermediary for marketing and advertising, clickworkers and gig workers have become human resources, and online shoppers have become an important target group (e.g., Dörr & Lautermann, 2024). The previous analogous stakeholder relationships are being transferred to the virtual world, creating business opportunities for forms of co-creation and collaboration, but also uncertainties about changes in stakeholder relationships and corporate responsibilities (Barnett et al., 2020; Jurgens et al., 2016).

Corporate Digital Responsibility (CDR) is gaining traction as an offspring of CSR (e.g., Lobschat et al., 2021). However, as the theoretical understanding of CSR has continuously evolved in recent decades, scholars are also examining CSR in combination with digitalisation (for CSR reviews, see Aslaksen et al., 2021; Carroll, 2021; Kumar et al., 2021; Latapí Agudelo et al., 2019; Rodriguez-Gomez et al., 2020). This raises the question of how CDR and CSR relate to each other: How does the digital dimension relate to environmental and social dimensions typically addressed by CSR? (Re)Setting conceptual boundaries is essential to avoid theoretical confusion, hence, in Chapter 3, the

question is posed: What are the essential differences between research on CDR and research that lies at the intersection of CSR and digitalisation? We conducted a systematic literature review and compared CDR and on CSR intersection digitalisation. Review results indicate that CDR primarily links information systems with business ethics, debating mostly theoretically on topics like integrating AI, the challenges of data security and gaining digital trust. By contrast, CSR is devoted more closely to the management community, examining empirical questions such as the measurement of performance, the interaction of sustainability and digitalisation, and the effects of CSR on digital communication. The stakeholder approach is pivotal in both research strands. At the same time, the measurement of policies and actual practices could be more differentiated (i.e., to sharpen the understanding between 'walking' and 'talking'). Further overlaps and research paths are outlined in the chapter.

1.2 Walk the Talk: Whistleblowing as a Signal for Credible CSR

Employees are the main stakeholders in implementing an ethical culture and supporting the company's CSR endeavours. Moreover, they are high in terms of stakeholder proximity (i.e., spatial nearness of the employee to the company), enabling them to identify and distinguish walking from talking, symbolic from substantial CSR actions. The latter describes actual, tangible, and measurable changes in business operations that require a firm's resources (Schons & Steinmeier, 2016). While employees' reaction to symbolic CSR actions is insignificant, employees have been shown to reward substantive CSR actions with, for instance, stronger firm commitment, which also positively affects firms financial performance (Aguilera et al., 2007; Schons & Steinmeier, 2016; H. C. Wang et al., 2009). Credible CSR actions can strengthen the relationship between employees and companies. As such, employees may provide feedback on whether management is complying with CSR and sensitise them to upcoming social responsibilities (De Roeck & Maon, 2018; Lin et al., 2022; Van Der Merwe & Al Achkar, 2022). By contrast, not walking the talk bears the risk of being perceived as machinewashing[2] or ethical washing (Bernini et al., 2024; Elliott & Copilah-Ali, 2024a). Such inconsistent, unauthentic, or misleading actions are often summarised in the concept

[2] "Machinewashing is defined as a business strategy for the ethical use of AI and algorithm-based systems, based on misleading behaviour affecting reporting (omitted or misleading information provided by words and images) and/or action (the underlying algorithm of AI) directed at various critical stakeholders to gain their acceptance" (Bernini et al., 2024: 329).

of Corporate Social Irresponsibility (CSIR) or Corporate Hypocrisy (e.g., Lin-Hi & Müller, 2013; Wagner et al., 2009)[3].

One such substantial CSR action is the implementation of whistleblowing mechanisms to call upon employees' responsibility to report observed wrongdoing. Whistleblowing is generally defined as „the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action" (Near & Miceli, 1985:4). As outlined, whistleblowing can be regarded as internalising CSR values such as transparency and accountability. Institutionalising an ethical culture needs to be backed with mechanisms that encourage reporting and protect employees from retaliation. After all, an experiment showed that participants were more likely to internally report fraud given a CSR condition than a non-CSR condition (Brink et al., 2018).

Whistleblowing can be divided into external and internal reporting. Different scandals have brought attention to external whistleblowing – the reporting of organisational wrongdoing to an outside authority. Examples include the NSA Surveillance Scandal (2013) by Snowden and, more recently, the Wirecard Scandal of 2020[4]. By contrast, internal whistleblowing can be referred to as the (anonymous) reporting of peers' misbehaviour, like fraud or norm violations, within the organisation. This allows the responsible manager to take appropriate steps to stop wrongdoing and introduce preventive measures before it can cause any further damage.

As the decision to report organisational misconduct is complex and controversial, research has been quite diverse in addressing the many facets of whistleblowing. Research in business ethics, management, and organisational behaviour has made important contributions by modelling whistleblowing as an ethical dilemma, describing how individuals act when observing wrongdoing in organisations (T. M. Jones, 1991; O'Fallon & Butterfield, 2005; Treviño, 1986). Other studies captured stages in the whistleblowing

[3] Corporate hypocrisy occurs when statements made and actual behaviour contradict each other, defined as "the belief that a firm claims to be something that it is not" (Wagner et al., 2009: 79). CSIR is defined as "corporate actions that result in (potential) disadvantages and/or harm to other actors" (Lin-Hi & Müller, 2013: 1932).

[4] NSA Surveillance Scandal (2013), where Edward Snowden, a former contractor for the U.S. National Security Agency (NSA), exposed extensive global surveillance programs conducted by the NSA and its allies; Wirecard-Scandal (2020), where a former head of legal at Wirecard Asia raised concerns and detected one of the largest financial frauds in German post-war history. Other prominent external whistleblowing cases are the Diesel (emissions) scandal in 2015 and the Facebook-Cambridge-Analytica Scandal in 2018.

decision-making process (Dozier & Miceli, 1985; Miceli & Near, 1992), while another literature strand examined how potential whistleblowers are influenced by individual characteristics such as a moral personality (Liyanarachchi & Newdick, 2009a; Vadera et al., 2009), or situational factors, such as the threat of work-related retaliation (Cassematis & Wortley, 2013; Mesmer-Magnus & Viswesvaran, 2005).

However, the picture of whistleblowing has changed, and a whistleblower is no longer solely regarded as a courageous act of an activist. Legislation (e.g., Sarbanes-Oxley Act, EU Whistleblower Protection Directive[5]) and recognition of organisations that internal whistleblowing is a valuable tool aim to create an environment to encourage whistleblowing, like providing whistleblowing hotlines or protection from retaliation (Vandekerckhove, Brown, & Tsahurid, 2014). Recently, researchers have pointed to the importance of the social environment for whistleblowers in the decision to report observed misconduct. As such, potential whistleblowers are influenced by peers' (perceived) expectations and potential social sanctions (Anvari et al., 2019; Lewis, 2022; Teo & Caspersz, 2011). This aligns with social influence research, which states that individuals are influenced by their immediate peer group's decisions, beliefs and attitudes (Abrams et al., 1990; Moore & Gino, 2013), and the social information processing perspective, which stresses that employees interpret cues from their work environment (Salancik & Pfeffer, 1978).

Chapter 4 precisely tackles this issue by reviewing and synthesising the increased but still fragmented literature on how social response influences an employee's whistleblowing behaviour in organisations, particularly co-workers. After all, whistleblowing tends to be polarised: while some see whistleblowers as heroes in the fight against crime, others see them as traitors to colleagues (Gagnon & Perron, 2020; Olesen, 2019). Hence, whistleblowers might weigh their loyalty towards the organisation and their co-workers. Seven thematical clusters are identified and merged into a framework. Five of these clusters are grouped into peer factors that influence a whistleblower, namely: (I) peer involvement in wrongdoing and whistleblowing situations; (II) allegiance to peers and the organisation; (III) behavioural prescription by

[5] European Sustainability Reporting Standards (ESRS), subsection G1 (business conduct), declares that corporations need to report the protection they provide for whistleblowers. Further, the Whistleblower Directive (EU) 2019/1937 sets a minimum standard for all EU countries, which needs to be transposed into national law, implying that companies must provide, for instance, whistleblower hotline, protection and training

peer; (IV) relationship and experiences with peers; and (V) fear of consequences from peers. Moreover, two clusters are attributed to adverse consequences that whistleblowers perceive in the aftermath: (VI) adverse perception that peers have concerning whistleblowers, comprising stigmatisation and likeability; and (VII) adverse actions that peers undertake, such as unofficial reprisals, ostracism, bullying, and the relationship with and social support for whistleblowers. After all, the review concludes that social expectations play a pivotal role for a potential whistleblower.

1.3 Behavioural Economic Approach to Emerging Issues on Social Norms

CSR and social expectation can be mutually reinforcing (see Figure 1). For one, companies are shaped by social expectations, consistent with the stakeholder approach, and aligned with the social contract theory, stating that society grants companies legitimacy as long as they serve society's interest (Donaldson & Dunfee, 1999; Heal, 2005). On the other hand, companies actively shape expectations, for instance, creating desires in terms of consumption. Also, ethical leadership and managers may signal which behaviour is appropriate in the work environment, influencing underlying normative expectations. Meanwhile, the investigations on the influence of social norms remain ongoing. Following Bicchieri (2017), social norms can be understood as "a rule of behavior such that individuals prefer to conform to it on condition that they believe that (a) most people in their reference network conform to it (empirical expectation), and (b) that most people in their reference network believe that ought to conform to it (normative expectation)" (Bicchieri, 2017: 35). As applicable from the previous chapter, organisational supporting whistleblowing is just one side of the coin. The social environment and peer factors are also crucial but may deviate from organisational norms or the CSR norms claimed by companies.

A potential whistleblower seldom knows the distributions of how others view whistleblowing but instead makes assumptions about the beliefs of others. Therefore, studying social norms in whistleblowing has the advantage of advancing both the current trends of social norms literature (e.g., identifying misperceptions, multiple expectations and the influence of norm interventions) and provides a viable context advancing the understanding of the micro-foundations of whistleblowing behaviour. While the previously introduced studies have taken a managerial perspective, another central part

of this dissertation is the behavioural economic approach to address the topic of social norms and whistleblowing. Before progressing thematically, the method and perspective of behaviour economic experiments will be outlined.

Behavioural economics includes psychological insights into economics, and a standard method is to conduct experiments. More concretely, the following studies conduct incentivised laboratory or online experiments for several reasons, also outlined at the end of Chapter 4. Laboratory experiments enable tight control over confounding variables and the decision environment (Hauser et al., 2017), providing a suitable empirical test environment for examining and identifying causal relationships derived from theories. Hence, laboratory experiments are characterised by a high internal validity but are often criticised for their limited external validity. However, whistleblowing and determinates that influence potential whistleblowers can hardly be observed in the field, and additionally, both witnessing and reporting wrongdoing tend to be rare and are often confounded by other factors. Moreover, interviews and surveys tend to ask about whistleblowing intention; however, actual whistleblowing behaviour may have different predictors (Mesmer-Magnus & Viswesvaran, 2005). As a consequence, the experimental approach is used in the following studies.

In Chapter 5, we examine misperceptions of normative expectations and the correction of them in the context of whistleblowing. Therefore, we conducted a survey on Prolific and an incentivised laboratory experiment. The results reveal that more than three-quarters of the individuals hold the personal normative belief that whistleblowing is the appropriate behaviour. However, almost half of the participants stated that they believe others believe whistleblowing is inappropriate (=normative expectations). While both personal normative belief and normative expectations predict whistleblowing behaviour in our experiment, individuals who are aligned – personally support whistleblowing and believe that others share their views – are more likely to report misconduct. In a second step, we introduced a social information intervention that reveals the true distribution of peer support from previous sessions, which affects subgroups differently. While it increases whistleblowing behaviour among individuals who already personally favour reporting misconduct, there is no effect among those who are personally resistant to it.

In Chapter 6, we focus on how multiple normative expectations about two mutually exclusive behaviours – whistleblowing and staying silent – individually and jointly influence whistleblowing behaviour. We designed a similar incentivised experiment as in Chapter 5 and conducted it on Prolific. The results indicate that normative expectation of whistleblowing being appropriate increases the reporting behaviour. Similarly, holding normative expectations for staying silent decreases whistleblowing behaviour. Moreover, this effect reinforces the behaviour of a specific subgroup: for those who believe that the majority supports whistleblowing, the reporting probability increases substantially when they simultaneously expect that staying silent is deemed inappropriate. Finally, we tested whether a social information intervention that contains a message in alignment with the majority norm elicited in the baseline treatment about either the appropriateness of whistleblowing or the inappropriateness of remaining silent or providing both messages influences reporting decisions differently. The findings show that the message about normative dimensions combined or only about the inappropriateness of staying silent significantly increases whistleblowing compared to the (no information) baseline and the message about whistleblowing appropriateness alone.

Chapter 7 investigates the methodical robustness of a norm elicitation measurement. More preciously, the commonly used two-step norm elicitation procedure is tested for its design features. The online experiment conducted on Prolific consists of a dictator game. It varies three design features, namely i) the time of elicitation (before vs after the behaviour), ii) incentivising vs not incentivising a question about normative expectations, and iii) questioning subjects on their beliefs about the action of interest alone or combined with an alternative action. A pretest reveals that applying role uncertainty in this context does not alter beliefs and behaviours compared to a baseline treatment without it. Subsequently, three treatments are implemented and reveal that – contrary to some previous results – the elicitation time does not alter the money-split decision. However, incentivising the question about normative expectations significantly increases the accuracy of answers (i.e., the correct estimation of the majority norm). Finally, asking about a fair share and an unfair share instead of only about fair sharing does not alter personal normative beliefs or normative expectations. However, it increases the empirical expectations that other dictators have provided a fair split.

Table 1 on the next page provides an overview of the studies (Chapters) in this dissertation, summarising the respective research questions, the primary contributions and the methods used. Table 2 shows the co-authors involved in the respective studies, as well as their shares and contributions. In addition, the current publication status and the presentations given for the respective study are shown (as of the beginning of April).

Table 1: Summary and Overview of the Studies in the Chapters

Research Objective	Contribution (limited to three main takeaways)	Perspectives	Method & Sample
Chapter 2: **From Digital Drift to Sustainable Direction – Reviewing the Interrelation of Corporate Sustainability and Digitalisation** Authors: Lena Epp (20%), Dörte Foit (20%), Tetiana Lutsenko (20%), Sabrina Plaß (20%), Thorben Scholz (20%)			
What do we (need to) know about how corporate sustainability interrelates with digitalisation?	1. Efficiency-driven perspectives dominate, focusing on ecological and economic benefits 2. Digitalisation is primarily framed as an enabler, not an independent goal of sustainability 3. strategic alignment is crucial for integrating digitalisation and sustainability within corporate practices	Corporate Sustainability (CS); Digitalisation	• Systematic Literature Review • A total of 3222 articles identified, after screening and selection process were 74 studies analysed in detail
Chapter 3: **Navigating Responsibility in the Digital Age: A Systematic Literature Review Comparing Corporate Digital Responsibility and Corporate Social Responsibility** Authors: Eva A. Jakob (35%) & Sabrina Plaß (65%)			
What are the essential differences between research on CDR and research that lies at the intersection of CSR and digitalisation?	1. Comparative analysis delineating the boundaries between CDR and CSR in the digital age by clarifying distinctions and commonalities 2. Advocates for a closer integration of the two fields 3. Highlights the need for more actionable, measurable frameworks	Corporate Social Responsibility (CSR); Corporate Digital Responsibility (CDR), Stakeholder	• Systematic Literature Review • After screening and selection 55 of 133 CDR articles and 75 out of 192 identified CSR and Digital* articles were analysed in detail and compared
Chapter 4: **Peers: Powerful or Negligible? A Systematic Review on Peer Factors and Internal Whistleblowing** Authors: Sabrina Loer (33%), Behnud Mir Djawadi (33%), Sabrina Plaß (33%)			
How are different peer factors related to internal whistleblowing?	Framework including: 1. Peer factors as antecedents & consequences of internal whistleblowing 2. Interaction of peer factors with other variables (i.e., moderating effects) 3. Future research agenda	Internal whistleblowing; Peers, Co-worker	• Systematic Literature Review • A total of 788 articles identified, after screening and selection process were 33 studies analysed in detail

Research Objective	Contribution (limited to three main takeaways)	Perspectives	Method & Sample
Chapter 5: **I don't believe that you believe what I believe: Experiment on Misperceptions of Social Norms and Whistleblowing** Authors: Sabrina Loer (33%), Behnud Mir Djawadi (33%), Sabrina Plaß (33%)			
How do potential whistleblowers personally think, what do they expect others to think about whistleblowing, and how is this relates to their behaviour? If there are misperceptions, does correcting the beliefs about others increase whistleblowing behaviour?	1. Misperceptions exist in the whistleblowing context 2. Personal normative beliefs and (misperceived) normative expectations influence whistleblowing 3. A social information intervention correcting misperceptions partly changes behaviour	Personal normative beliefs; Normative expectations; Internal whistleblowing	• Survey on Prolific (n=100), UK participants, currently employed, witnessed wrongdoing in the workplace • Incentivized experiment (BaER-Lab), 2 treatments (n=396)
Chapter 6: **Multiple Normative Expectations and Interventions – Experimental Evidence on Whistleblowing Behaviour** Authors: Sabrina Loer (33%), Behnud Mir Djawadi (33%), Sabrina Plaß (33%)			
To what extent do both normative expectations (regarding whistleblowing and remaining silent) relate to the whistleblowing decision? Does providing information about majority beliefs (individually and jointly) affect whistleblowing behaviour?	1. Normative expectations regarding the appropriateness of whistleblowing and staying silent are both individually and jointly related to whistleblowing behaviour 2. Distribution of normative expectations matters 3. Interventions highlighting silence as inappropriate significantly increases whistleblowing behaviour	Multiple normative expectations; Social information interventions; Internal whistleblowing	• Incentivized online experiment on Prolific: Baseline + 3 Treatments • Sample size of n=367, UK participants full or half time employed)
Chapter 7: **Testing Variations of the Two-Step Norm Elicitation Procedure** Author: Sabrina Plaß (100%)			
Do (and if, how do) variations in the two-step norm-elicitation method (e.g., Bicchieri & Xiao, 2009) affect subjects indicated normative expectations and behaviour?	1. Behaviour does not significantly differ when eliciting normative expectations (NE) before the task compared to the elicitation after the task 2. Incentivization significantly increases the accuracy of normative expectations 3. Asking about two behavioural alternatives increases empirical but not normative expectations	Personal normative beliefs; Normative expectations; Dictator game; Fairness	• Incentivized online experiment on Prolific: Pre-Test + Baseline + 3 Treatments • Sample size of n=639, UK participants

Table 2: Authors Contribution, Status of Publication and Conferences

From Digital Drift to Sustainable Direction – Reviewing the Interrelation of Corporate Sustainability Integrating and Digitalisation *by Lena Epp, Dörte Foit, Tetiana Lutsenko, Sabrina Plaß, Thorben Scholz (all 20%)*
• Idea and study development: jointly • Methodical implementation: jointly • Framework development & Write-up of paper: jointly
Publication & Conferences: • Submitted in February 2025: Conference of *British Academy of Management (BAM)* • (Planned) Submission in April 2025: *Journal of Strategic Management*
Navigating Responsibility in the Digital Age: A Systematic Literature Review Comparing Corporate Digital Responsibility and Corporate Social Responsibility *by Eva A. Jakob (35%), Sabrina Plaß (65%)*
• Idea and study development: jointly • Methodical implementation: mainly S. Plaß • Framework development: mainly E.A. Jakob • Write-up of paper: mainly S. Plaß
Publication & Conferences: • Previous version in *Academy of Management Proceedings* (Title: Corporate Digital Responsibility Needed? Digitalization Meets Corporate Social Responsibility) • Currently in revision (until April): *Schmalenbach Journal of Business Research* Previous version presented by: • Jakob & Plaß at CDR-Initiative Event by BMJV[6] 2019, Berlin • Plaß at *Stakeholder Summer School at Darden University*, Charlottesville, Virgina 2019 • Jakob at *Annual Meeting - Academy of Management* 2022, online
Peers: Powerful or Negligible? A Systematic Review on Peer Factors and Internal Whistleblowing *by Sabrina Loer (33.33%), Behnud Mir Djawadi (33.33%), Sabrina Plaß (33.33%)*
• Idea and study development: jointly • Theoretical classification: mainly by B. Mir Djawadi • Methodical implementation: mainly by S. Loer & S. Plaß • Framework development & Write-up of paper: jointly
Publication & Conferences: Published in C. Gabbioneta, M. Clemente, & R. Greenwood (Hrsg.), *Organizational Wrongdoing as the "Foundational" Grand Challenge: Consequences and Impact* (S. 73–100). Emerald Publishing Limited. *Research in the Sociology of Organizations, 85.* (VHB Ranking: B)

[6] BMJV: Bundesministerium der Justiz und für Verbraucherschutz

I don't believe that you believe what I believe: Experiment on Misperceptions of Social Norms and Whistleblowing

by Sabrina Loer (33.33%), Behnud Mir Djawadi (33.33%), Sabrina Plaß (33.33%)

- Idea and experimental design: jointly
- Programming of experiment in o-tree: mainly by B. Mir Djawadi
- Testing: jointly
- Data collection: jointly
- Preparation and formalization: mainly S. Plaß
- Analysis: mainly by S. Loer
- Write-up of paper: jointly

Publication & Conferences:

Currently in revision (until May): *European Economic Review*

- Presented by Plaß at SABE/ IAREP[7] Conference 2023 in Nice (France)
- Presented by Plaß at GfeW[8] 2023 in Erfurt
- Presented by Loer at SABE/ IAREP Conference 2024 in Dundee (Scottland)

Multiple Normative Expectations and Social Norm Interventions – Experimental Evidence on Whistleblowing Behaviour

by Sabrina Loer (33.33%), Behnud Mir Djawadi (33.33%), Sabrina Plaß (33.33%)

- Idea and experimental design: jointly
- Programming of experiment in o-tree: mainly by B. Mir Djawadi
- Testing: jointly
- Data collection: S. Loer & S. Plaß
- Preparation and formalization: mainly S. Plaß
- Analysis: mainly by S. Loer
- Write-up of paper: jointly

Publication & Conferences:

- Submitted in March 2025: *Journal of Management Science*
- Presented by Plaß at *SABE/ IAREP Conference* 2024 in Dundee (Scottland)
- Presented by Plaß at *GfeW* 2024 in Cologne

Testing Variations of the Two-Step Norm Elicitation Procedure

by Sabrina Plaß (100%)

Publication & Conferences:

- Submitted in March 2025 and currently under review: *Journal of Economic Behavior & Organization*
- Accepted for presentation at *SABE 2025 (in June)* in Trento (Italy)

[7] SABE/ IAREP: Society for the Advancement of Behavioral Economics (SABE) / International Association for Research in Economic Psychology (IAREP)
[8] GfeW: Gesellschaft für experiementelle Wirtschaftsforschung

CHAPTER 2 | From Digital Drift to Sustainable Direction – Reviewing the Interrelation of Corporate Sustainability and Digitalisation

Lena Epp Dörte Foit Tetiana Lutsenko Sabrina Plaß Thorben Scholz

The interplay between corporate sustainability and digitalisation is increasingly recognised as dual transformation, yet its strategic integration remains underexplored. This study systematically reviews quantitative and qualitative research on the corporate-level intersection of sustainability and digitalisation. Thematic clustering of their impacts led to the following five key insights: (1) research lacks integrated, multi-level approaches; (2) digitalisation is primarily framed as an enabler, rather than an independent goal of sustainability; (3) strategic alignment is crucial for integrating digitalisation and sustainability within corporate practices; (4) efficiency-driven perspectives dominate, focusing on ecological and economic benefits; (5) methodologies and topics are diversifying, with growing emphasis on social and governance dimensions. By systematically structuring existing knowledge and critically assessing key challenges and opportunities, this study provides a comprehensive foundation for future research on the strategic integration of dual transformation.

Keywords: *Sustainability, Digitalisation, Systematic Literature Review, Strategy, Transformation*

2.1 Introduction

Digitalisation[9] and sustainability are two of the most powerful transformation drivers of our time (Flyverbom et al., 2019). Yet, too often, they unfold in parallel rather than in sync. Digitalisation risks becoming a digital drift when pursued in isolation - an adoption of digital technologies without a clear sustainability trajectory. By contrast, sustainability efforts that neglect digital capabilities may struggle to scale, remaining reactive rather than transformative. Not choosing one path over the other but aligning both to create a strategic direction is often referred to as dual or twin transformation (Barth et al., 2023; Epp et al., 2024). While some companies are beginning to explore these synergies, research remains fragmented, leaving open questions about how businesses can strategically align both forces for long-term impact (Broccardo et al., 2023).

To address challenges and potentials, digitalisation needs to be strategically aligned with the long-term objectives of sustainable development (Brenner & Hartl, 2021; Legner et al., 2017a). From a holistic economic perspective, this implies the preservation of ecological, social and economic capital stocks as complements rather than substitutes (Daly, 1995, 1996; DesJardins, 2007). This perspective conceptualises an interconnected system in which economic activities are embedded and constrained by ecological and social boundaries to achieve long-term sustainable development (Bansal & Song, 2017; Giddings et al., 2002; Raworth, 2017). For corporations, sustainability refers to actions that (when aggregated) contribute to this comprehensive goal. However, the existing literature remains fragmented, with limited emphasis on the strategic integration of sustainability and digitalisation at the corporate level. Accordingly, we pose the question: *What do we (need to) know about how corporate sustainability interrelates with digitalisation?* This study follows a structured approach closely aligned to the classical systematic review methodology as outlined by Denyer and Tranfield (2009).

Before analysing the interrelation between digitalisation and sustainability, each of these transformation drivers need to be examined individually. Management scholars have proposed partly overlapping conceptual approaches to describe firms' efforts

[9] Terms like digitisation (or 'datafication'), digitalisation and digital transformation vary across and within disciplines. In this paper, we use the term digitalisation to describe how organisations and users generate and apply digitised data, utilise digital technologies, including artificial intelligence (AI), and drive innovation in business models, often disrupting markets (including through digital transformation, Brenner & Hartl, 2021; Legner et al., 2017).

towards sustainability. These comprise, but are not limited to, *Corporate Social Responsibility* (CSR, e.g., Garriga & Mele, 2004), which focuses on responsibilities and normative conduct; *Corporate Social Performance* (CSP, e.g., Waddock & Graves, 1997), an instrumental perspective examining the outcomes of social business practices; the *Triple Bottom Line* (TBL, Elkington, 1998), which adopts a long-term perspective on the interplay between economic, environmental and social dimensions; and *Stakeholder Theory*, emphasising the integration of those affected by or affecting the firm's activities (e.g., Parmar et al., 2010). Among these conceptual approaches, *Corporate Sustainability* (CS) has emerged as a widely used umbrella term that encompasses related concepts such as CSR and CSP (Montiel, 2008). However, its definition and concept remain fragmented and ambiguous, with no standardisation across disciplines[10] (Garriga & Mele, 2004; Van Marrewijk, 2003; Van Zanten & Van Tulder, 2021). In this study, we adopt CS as our focal concept describing organisations' contributions to long-term sustainable development through incorporating corresponding endeavours into corporate strategies and practices (Denyer & Tranfield, 2009). Scholars emphasise that effective CS requires strategies and practices tailored to a company's core business (Garriga & Mele, 2004; Van Marrewijk, 2003; Van Zanten & Van Tulder, 2021). Recent CS research differentiates between examining CS outcomes, for example, how specific actions and strategies contribute to sustainable development (e.g., Van Zanten & Van Tulder, 2021), and the study of CS approaches and knowledge, for example, how managers require competencies, implement and promote sustainability (strategies) in their organisation (Strand, 2023). Reviews on CS have mainly focused on specific domains like supply chain management (Chauhan et al., 2023; Strand, 2023), reporting/ auditing (Ferrante et al., 2024; Hina et al., 2022) and circular economy (Ferrante et al., 2024; Hina et al., 2022).

Only recently the intersection of sustainability and digitalisation has sparked scholarly interest (e.g., Del Río Castro et al., 2021; Isensee et al., 2020). Two main terms have been introduced in this context. First, *Digital Sustainability* (Del Río Castro, Gonzalez Fernandez, et al., 2021; Isensee et al., 2020) describes the synergy effect of digitalisation enabling the achievement of sustainability goals, such as circular economy and efficiency gains. Second, *Sustainable Digitalisation* refers to applying sustainability criteria and practices in creating digital environments while mitigating potential

[10] For an overview you may refer to Meurer (2020) who outlines 33 existing definitions of Corporate Sustainability or to Amini and Bienstock (2014) who derived a CS framework based on a focused review.

(ecological or social) rebound effects and managing changes in stakeholder dynamics (e.g., Barth et al., 2023; Epp et al., 2024; Feroz et al., 2023; Guandalini, 2022). Much as these concepts indicate a growing body of research, a critical review is essential to synthesise existing knowledge, identify gaps, and provide a coherent understanding and shared terminology for both scholars and practitioners.

This study contributes to management literature by systematically reviewing the intersection of CS and digitalisation, addressing the fragmentation in existing research and exploring how strategic and sustainability dimensions are conceptualised. It provides insights into dominant research perspectives and identifies prevailing thematic and methodological orientations and areas that require further investigation. Additionally, the study critically examines how digitalisation is positioned within sustainability discourses and assesses the role of corporate strategy in aligning these transformations. By synthesising existing knowledge, this review offers a structured foundation for future research and supports a more integrated perspective on corporate sustainability and digitalisation.

2.2 Research Process and Method

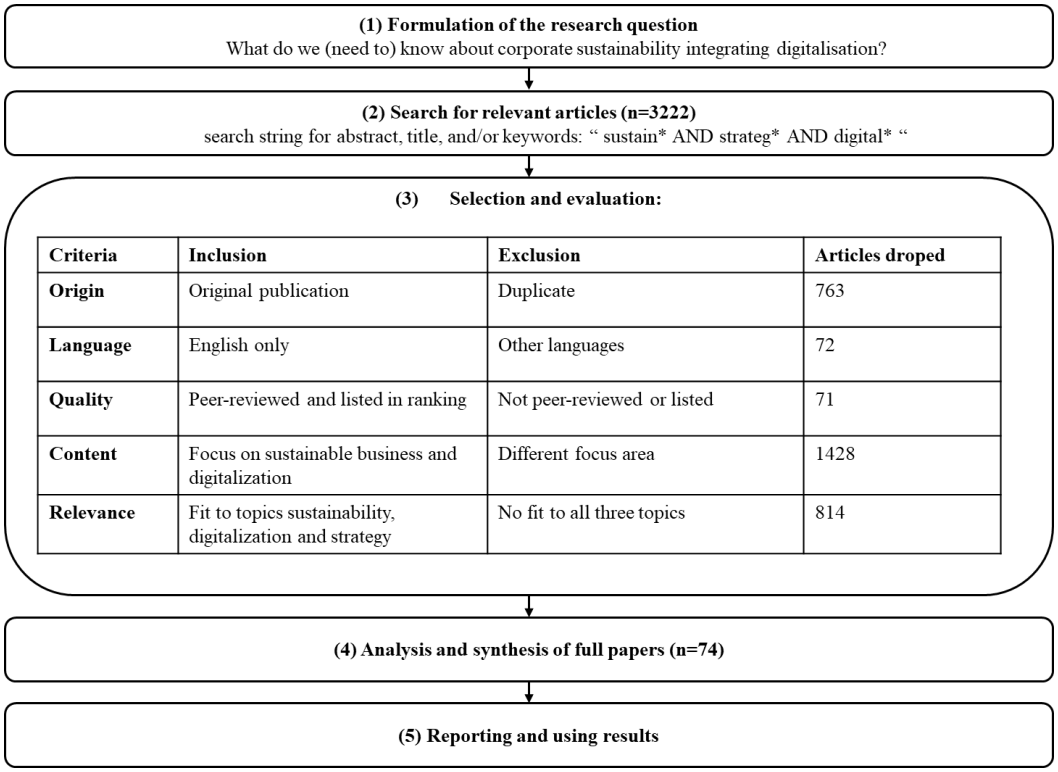
This study applies a systematic literature review and critically reflects on the current state of research regarding CS and its integration with digitalisation, providing insights for management research at the intersection of these megatrends. Our review follows the five step-approach by Denyer and Tranfield (2009). According to the classification by Krlev et al. (2025)[11], our approach primarily qualifies as a taking stock review, as it systematically categorises fragmented research to provide a structured overview.

Figure 1 provides detailed information about the search, selection and evaluation process, starting with *(1) formulation of the research question*. The second step, *(2) search for relevant articles*, refers to the general search process conducted between June and August 2023 using the databases Scopus, Web of Science, and EBSCOhost. By applying the terms *sustain** and *digital** in the search string, we ensured the inclusion of all relevant corresponding terms, avoiding a narrow focus on specific concepts. Further, we included *strateg** in the search string to link findings to the strategic dimension of

[11] Recently, Krlev et al. (2025) have provided a systematic overview of ten different types of literature reviews locating review purposes on directional space for reviewing (degree of substantiveness and reflexiveness).

corporate sustainability rather than solely focusing on individual aspects, for example, product innovations or operational procedures[12].

Figure 1: Search and Selection Process



Criteria	Inclusion	Exclusion	Articles droped
Origin	Original publication	Duplicate	763
Language	English only	Other languages	72
Quality	Peer-reviewed and listed in ranking	Not peer-reviewed or listed	71
Content	Focus on sustainable business and digitalization	Different focus area	1428
Relevance	Fit to topics sustainability, digitalization and strategy	No fit to all three topics	814

In the third step, *(3) selection and evaluation*, we consecutively performed a formal screening, reviewing the titles and abstracts, before screening each paper in full. To ensure quality, we primarily considered papers published in journals listed in the VHB ranking (independent of their individual rating). Nevertheless, 18 articles from non-VHB-ranked journals were included if they fit the research topics and demonstrated the potential to considerably contribute to answering our research question. For example, Demir et al. (2023) provide a readiness and maturity model for sustainable supply chains, while Ghobakhloo, Asadi, et al. (2023) survey 207 multinational firms on the impact of intelligent automation on sustainability performance.

The retained papers were screened twice by a group of five reviewers, who worked independently, focusing on the papers' title and abstract. This process was repeated during the full paper screening step. Papers were excluded if their content considerably diverged

[12] To ensure this, we further set a focus on relevant categories in the databases used: in Web of Science, we searched categories Environmental Studies, Environmental Sciences, Management, Business, and Economics; in Scopus the categories Business, Management and Accounting, Economics, Econometrics, and Finance; in EBSCOhost the database Business Source Complete.

from the research scope, particularly those focusing on sectors unrelated to corporate sustainability and digitalisation. Accordingly, we excluded studies that examined contexts beyond companies, such as education (Hashim et al., 2022), digitalised agriculture (Biró & Szalmáné Csete, 2022; Prihadyanti & Aziz, 2023), or underground hydrocarbon storage (Zhang et al. 2020). Articles missing at least one of the three core components of the search string (e.g., missing/weak focus on sustainability, digitalisation, or no reference to corporate/industry) were excluded. Inclusion was based on the contextual alignment of sustainability, digitalisation and strategy with the research question.

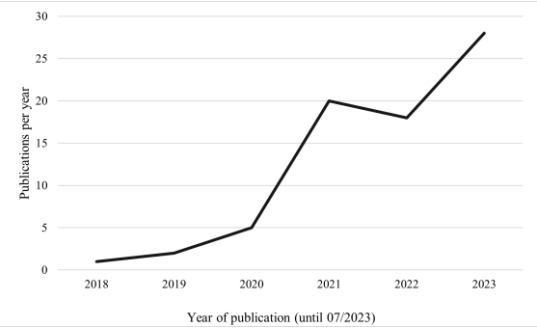
In *(4) analysis and synthesis*, the remaining 74 articles were analysed in two rounds of independent reviewing, whereby all the information from the articles was screened in full. The result is the systematic analysis of insights from research on CS and digitalisation. Step five, *(5) reporting and using results*, is presented in Sections 4 and 5.

2.3 Descriptive Review

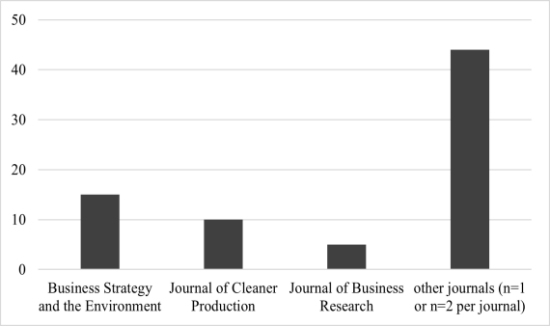
This section provides insights into publication trends, thematic focus areas, and methodological approaches of the 74 retained articles (for an overview see Appendix A of Chapter 2, p. 203). The most frequently represented journals include *Business Strategy and the Environment* (n=15), *Journal of Cleaner Production* (n=10), and *Journal of Business Research* (n=5), with other journals contributing one or two publications each, indicating a wide dispersion across different academic outlets (Figure 2b). The number of articles has increased significantly since 2021, reflecting a growing academic interest in corporate sustainability and digitalisation (Figure 2a).

Figure 2: Year of Publication and Most Common Journals of the 74 Sampled Articles

a) Year of Publication over Time



b) Journals and Number of Articles



A notable pattern in the dataset is the fragmented yet interdisciplinary nature of the research field, evident in the diversity of study scopes and themes. A large proportion of studies adopt a global perspective (n=33), particularly conceptual papers and literature

reviews without regional or industry-specific constraints. Other studies focus on specific countries (n=25), with China, India, and Italy the most frequently represented. Additionally, 16 articles investigate specific industries, primarily manufacturing or logistics, while only three articles analyse individual company cases. In terms of thematic focus, Industry 4.0 (n=25) and circular economy (n=19) are the most frequent topics, with 12 articles explicitly examining their interrelation. Figure 3 presents the articles' keywords, highlighting thematic areas with at least three entries related to sustainability (e.g., sustainable manufacturing, circular economy), digitalisation (e.g., Industry 4.0, digital transformation), and related terms (e.g., economic performance or stakeholder engagement). It is worth noting that some terms, like strategy or supply chain, are not mentioned as keywords even though they are present in many articles.

Figure 3: Word Cloud of Literature Keywords

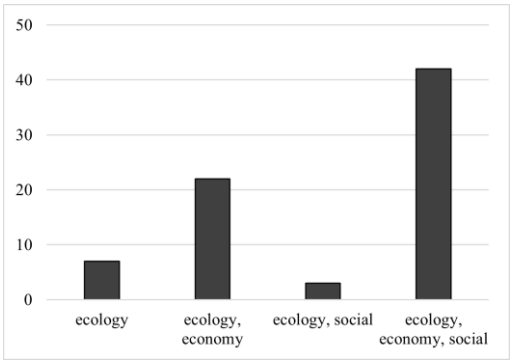
The conceptualisation of sustainability varies across studies, with most addressing all three sustainability dimensions (n=42; Figure 4a). However, the ecological dimension receives particular emphasis. Regarding the relationship between sustainability and digitalisation, 53 articles focus on digital sustainability (Figure 4b). In contrast, only one article addresses sustainable digitalisation, and 19 mention both directional relationships.

Regarding data collection methods (Figure 5a), empirical (n=35) and non-empirical (n=39) approaches are almost evenly distributed. The most frequently used methods are surveys (n=22) and various types of literature reviews (n=23), including 18 systematic literature reviews, four literature reviews and one integrated literature review. Furthermore, ten more articles conducted unstructured literature analyses. Additionally, out of the eight articles that employed a mixed-methods approach, three combined a

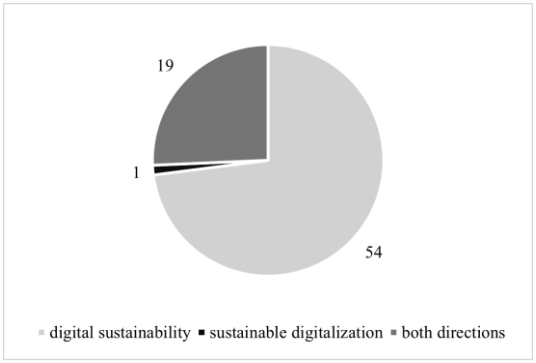
literature analysis with expert opinions, three with surveys, and four with case studies. Moreover, eight articles conducted case studies, two used panel data and one interviews.

Figure 4: Thematical Coverage

a) Sustainability Dimensions covered



b) Relationship between Sustainability and Digitalisation

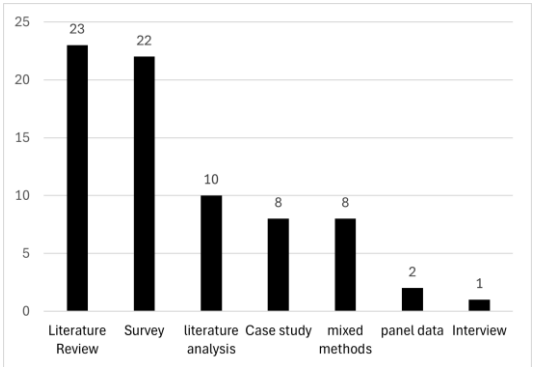


c) Thematic Focus over Publication Time

Year	Thematic Foci
2018–2020	Industry 4.0, circular economy, supply chains; initial conceptual discussions on digitalization & sustainability, limited empirical research.
2021	Digital transformation emerges as a key topic; increasing focus on business models, innovation, and strategy; Industry 4.0 and circular economy remain dominant.
2022	Growing attention to AI, Big Data, intelligent automation; early discussions on social sustainability & governance; supply chain sustainability remains relevant.
2023	Expansion to sustainable digitalization, twin transition, Net-Zero Manufacturing; increasing studies on strategic perspectives, stakeholder engagement, and organizational capabilities.

Figure 5: Methods of Data Collection

a) Methods of Data Collection of retained Articles



b) Methodical Focus over Publication Time

Year	Methodic Foci (of Data Collection)
2018–2020	Literature reviews, conceptual studies; limited empirical analyses.
2021	More surveys, case studies, early mixed-methods approaches (expert opinions & empirical data).
2022	Increasing use of quantitative models (e.g., SEM, regression); growing reliance on secondary firm data.
2023	Rise in Delphi studies, engaged scholarship, systematic combining; stronger integration of qualitative & longitudinal methods.

Our analysis shows that thematic and methodical foci have evolved over time (Figure 4c & 5b). Between 2018 and 2020, studies primarily focused on conceptualising the relationship between digitalisation and sustainability, emphasising technological enablers such as Industry 4.0, smart manufacturing, and big data analytics. Literature reviews dominated the methodological landscape, laying the theoretical groundwork for

future empirical studies. From 2021 onwards, research expanded beyond technological drivers to include corporate sustainability perspectives, circular economy frameworks, and supply chain management, reflecting a broader scope of inquiry. The growing interest in the strategic and organisational dimensions of sustainability is mirrored in an increased use of survey-based studies and quantitative modelling approaches such as structural equation modelling (SEM) and regression analysis. Methodological shifts further intensified in 2022 and beyond, with a rising number of qualitative and mixed-methods studies. Case studies, expert interviews, and engaged scholarship approaches gained prominence, particularly in research focusing on corporate strategy, governance mechanisms, and sustainability transitions. This shift indicates the growing recognition that digitalisation and sustainability transformations cannot be understood through efficiency metrics alone, but require context-sensitive, interdisciplinary research designs.

2.4 Thematic Review Results

The following thematic analysis adopts an inductive approach. Initially, we coded all variables from 32 empirical articles according to the dominant thematic topic (Table 1). Subsequently, we derived five higher order thematical clusters based on content similarities[13]. The clusters depict the relationships between dependent and independent variables, as well as between potential mediators and moderators[14] (Table 2). This approach provides a comprehensive overview of the empirical findings in the research field. In addition to the quantitative studies, we assign qualitative studies to the most thematically similar cluster.

To ensure reliability, two researchers independently coded the variables and categorised them into clusters, refining the results through an iterative process. Although some thematical overlaps emerged, we coded and clustered each factor to the dominant theme to avoid redundancies. The ‘strategy’ topic often serves as a meta-component that complements the primary theme of a variable. To acknowledge this specific overlap, we marked such variables with an asterisk (*). In the subsequent paragraphs (4.1 - 4.5), each thematic cluster is described and its analysis introduced with an impact map (Figure 6 -

[13] The procedure systematically groups the sequences of variables according to similar characteristics, aligning with established methodologies for identifying overarching thematic clusters in systematic literature reviews (e.g., Guandalini, 2022).
[14] Moderators appear only in a few cases (L. J. Zheng et al., 2023a) and are marked accordingly.

11), followed by a detailed profile (Table 3 - 7) that compiles key characteristics, to provide a structured foundation for the in-depth review.

Table 1: Assignment Criteria for Coding of Variables

Code	Topic	Assignment Criteria
DT	Digitalisation	Variables focusing on digital technologies, digital transformation, digital infrastructures, and digital tools. Includes elements like AI, big data, blockchain, digital strategies, and digital infrastructures.
EM	Economy	Variables addressing economic performance, profitability, cost-efficiency, market competitiveness, and financial outcomes.
EL	Ecology	Variables targeting environmental aspects, such as carbon reduction, resource efficiency, circular economy practices, and environmental sustainability.
SC	Social	Variables related to social sustainability, stakeholder relationships, employee engagement, customer collaboration, and broader social impacts.
ST	Strategy	Variables dealing with strategic planning, management approaches, leadership, strategic decision-making processes, and innovation.
SU	Sustainability (multidimensional)	Variables reflecting integrated sustainability aspects that encompass at least two or all three sustainability dimensions (economic, ecological, social).

Table 2: Thematic Clusters and Code Patterns

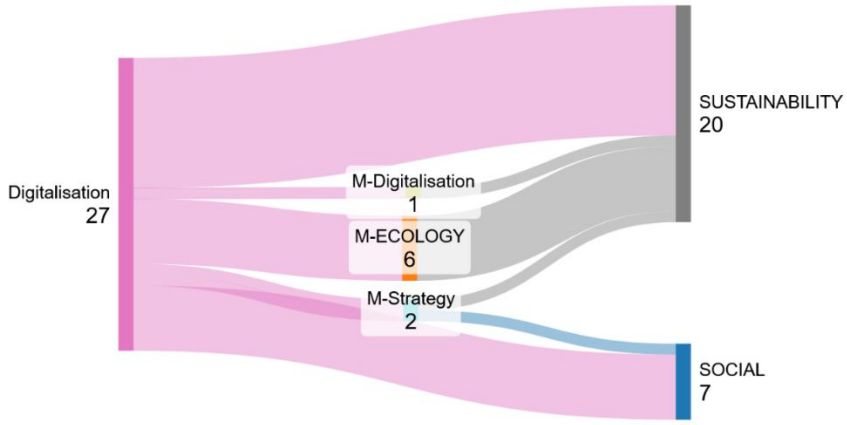
Cluster Title	Description	Code Patterns
Digitalisation as Enabler for Sustainability	Focuses on how digital technologies serve as drivers for achieving sustainability outcomes across multi-dimensional (and social) domains.	*DT → SU (/SC)*
(Digitally Driven) Economic Efficiency	Investigates how digital technologies improve economic efficiency through profitability, operational optimisation, and enhanced competitiveness.	*DT → EM (ST/SC)*
(Digitally Driven) Ecologic Efficiency	Focuses on how digital technologies contribute to ecological efficiency, particularly through resource optimisation, carbon reduction, and improved environmental performance.	*DT/ EL → EL*
Sustainability-Driven Business and Digitalisation	This cluster examines how sustainability considerations – including environmental, social, and multi-dimensional sustainability – serve as drivers for shaping business performance, digitalisation, and strategies.	*SU/SC/EL → EM/DT/ST*
Strategy as a Catalyst for Sustainable Development	This cluster focuses on the role of strategic approaches in driving sustainability-oriented transformations. Strategies serve as catalysts for improving social, multi-dimensional sustainability, and ecological outcomes.	*ST → SU/SC/EL*

Digitalisation as Enabler for Sustainability

This cluster comprises 27 relationships that appear in nine quantitative empirical studies. These studies explore how independent variables in the context of digitalisation act as a primary driver for sustainable performance and implementation of sustainability

(see Table 3). The Cluster Impact Map[15] (Figure 6) illustrates the relationships between digitalisation impact category, mediators, and sustainability outcomes. Most pathways show a direct positive influence, with ecology-related mediators the most prominent. Though social outcomes are less studied, they remain a relevant dimension in this cluster.

Figure 6: Cluster Impact Map – Digitalisation as Enabler for Sustainability



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right (more detailed in the table below). "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Table 3: Cluster Profile – Digitalisation as Enabler for Sustainability

Source	Independent Variable	Mediator [MODerator]	Dependent Variable	Code Pattern
Ardito, 2023	Smart Devices (+)		SUSTAINABLE INNOVATION * (SOCIAL AND ENVIRONMENT)	*DT → SU**
	Blockchain (+)			*DT → SU**
	Big Data Analytics (+)			*DT → SU**
	AI (+)			*DT → SU**
	Robots (+)			*DT → SU**
	High Speed Infrastructure (+)			*DT → SU**
Pînzaru et al., 2022	Digitalization (+)		SUSTAINABILITY PRACTICES	*DT → SU*
			ORGANIZATIONAL OUTCOMES (OF SUSTAINABILITY)	*DT → SU*
Broccardo et al., 2023	Digitalization (+)		SUSTAINABILITY IMPLEMENTATION AND RELATED PERFORMANCE	*DT → SU*
Belhadi et al., 2022	Industry 4.0 Capabilities (+)	Digital Business Transformation * (+)	SUSTAINABLE PERFORMANCE	*DT → DT* →SU*

[15] The Cluster Impact Maps illustrate only statistically tested relationships, regardless of the direction or significance of the impact.

		Organizational Ambidexterity (+)		*DT → ST →SU*
	Digital Business Transformation (+)			*DT → SU*
Xu et al., 2023	Digital Capability (+)	Eco-management Innovation* (+)		*DT → EL* →SU*
		Eco-process Innovation* (+)		*DT → EL* →SU*
		Eco-product Innovation* (+)		*DT → EL* →SU*
	Digital Strategy (+)	Eco-management Innovation* (+)		*DT → EL* →SU*
		Eco-process Innovation* (+)		*DT → EL* →SU*
		Eco-product Innovation* (+)		*DT → EL* →SU*
Zheng et al., 2023	Internal Digitalization (+)	[MOD: Ownership Diversification (+)]	CSR PERFORMANCE	*DT → SU*
	External Digitalization (+)			*DT → SU*
Al-Khatib, 2023	Industrial Internet of Things (+)	Supply Chain Visibility (+)	SOCIAL PERFORMANCE	*DT → ST →SC*
Ghobakhloo, Asadi, et al., 2023	Intelligent Automation Implementation (0)		SOCIAL SUSTAINABILITY PERFORMANCE	*DT →SC*
Ferreira et al., 2023	Cloud Computing (+)		SOCIAL SUSTAINABILITY	*DT →SC*
	Big Data Analytics (+)			*DT →SC*
	Robotics (+)			*DT →SC*
	AI (+)			*DT →SC*
	Blockchain (+)			*DT →SC*

Qualitative Insights (not statistically tested[16])

Source	Impact Factor	Outcome Factor
Santarius & Wagner, 2023	ICT	Sustainability
Guandalini, 2022	Digitalization	
Acciarini et al., 2022		Sustainable Innovative Business Models
George et al., 2021	Digital Solutions for Managerial Problems	Sustainable Firm Development
Islam et al., 2022	Digital Performance	Smart Sustainable Business Growth Model
Ghobakhloo, Iranmanesh, et al., 2023	Industry 5.0	Sustainable Industrial Transformation
Bag et al., 2021	Industry 4.0 Enabler	Sustainable Supply Chain
Patil et al., 2023	Big Data-Industry 4.0 Interaction	
Pan & Nishant, 2023	AI	SDG
Ching et al., 2022	Industry 4.0 Applications	Sustainable Manufacturing

[16] Our classification distinguishes between empirical and non-empirical studies, further subdividing empirical research into quantitative and qualitative categories. However, for the analysis of impact pathways, we focused on statistically tested relationships. Consequently, some studies categorised under 'qualitative insights' employ quantitative data collection methods but do not present statistically tested variable relationships (Feroz et al., 2023; Neligan et al., 2023).

Kumar et al., 2022	Industry 4.0	
Jamwal et al., 2022	Deep Learning	
Agrawal, Majumdar, et al., 2023	Integration of AI	
Kumar et al., 2021	Strategic Factors for Application of Big Data Analytics	
Ghobakhloo et al., 2021	Industry 4.0 Functions	Sustainable Innovation in Manufacturing
Rejeb & Rejeb, 2020	Blockchain	Supply Chain Management
Sahu et al., 2023		
Contini et al., 2023	Monitoring through KPI and Sustainable Digital Twin	Sustainability Performance in Production
Sætra, 2023	AI ESG Protocol	

Note: The direction and quality of effects is indicated as follows: (+) positive effect; (-) negative effect; (0) not significant.

Digitalisation plays a pivotal role in embedding sustainability into business operations and fostering innovation that aligns with environmental, social, and economic objectives. Broccardo et al. (2023) demonstrate that digital technologies enable the more effective integration of sustainability into business operations, leading to improved sustainability outcomes. However, the adoption of digital tools for sustainability is often incremental, with many firms regarding them as supportive mechanisms rather than disruptive agents of change. This view is echoed by Pînzaru et al. (2022), who find that digitalisation promotes the adoption of SUSTAINABILITY PRACTICES, such as setting key performance indicators, fostering a sustainability-oriented culture, and encouraging employee participation. These findings highlight the instrumental dimension of digitalisation, where technologies are leveraged to embed sustainability into everyday operations.

Beyond implementation, studies point to performance improvements resulting from digital-driven sustainability efforts. Xu et al. (2023)highlight eco-innovation as a key mechanism linking digital transformation to SUSTAINABLE PERFORMANCE and demonstrating that firms with strong digital capabilities and a focus on eco-innovation are more likely to achieve enhanced environmental, economic, and social outcomes. Similarly, Zheng et al. (2023) reveal that internal and external digitalisation positively influence CSR PERFORMANCE, although the magnitude of this effect is shaped by governance factors such as ownership diversification. Ardito (2023) further examines the link between digitalisation and sustainability performance by investigating how technologies like smart devices or blockchain influence firms' SOCIAL AND ENVIRONMENTAL INNOVATIONS. Blockchain is notably associated with social innovations,

whereas smart devices are linked to environmental innovations. Despite the generally positive influence of digital technologies on sustainable innovation, the effectiveness of this relationship can vary depending on organisational context and technology type. Belhadi et al. (2022) reinforce this narrative by highlighting the importance of organisational ambidexterity in translating digital capabilities into SUSTAINABLE PERFORMANCE improvements.

Extending this performance focus, Al-Khatib (2023) identifies a positive relationship between the adoption of industrial Internet of Things (iIoT) and SOCIAL PERFORMANCE, mediated by improved supply chain visibility, while Ghobakhloo, Asadi, et al. (2023) find no significant direct impact of intelligent automation implementation on SOCIAL SUSTAINABILITY PERFORMANCE. This contrast highlights that not all advanced technologies uniformly enhance social sustainability outcomes, suggesting that technology-specific capabilities and integration processes are critical determinants. Ferreira et al. (2023) support this notion by demonstrating that, among the five analysed technologies, cloud computing has the strongest impact on SOCIAL SUSTAINABILITY, while blockchain has the weakest.

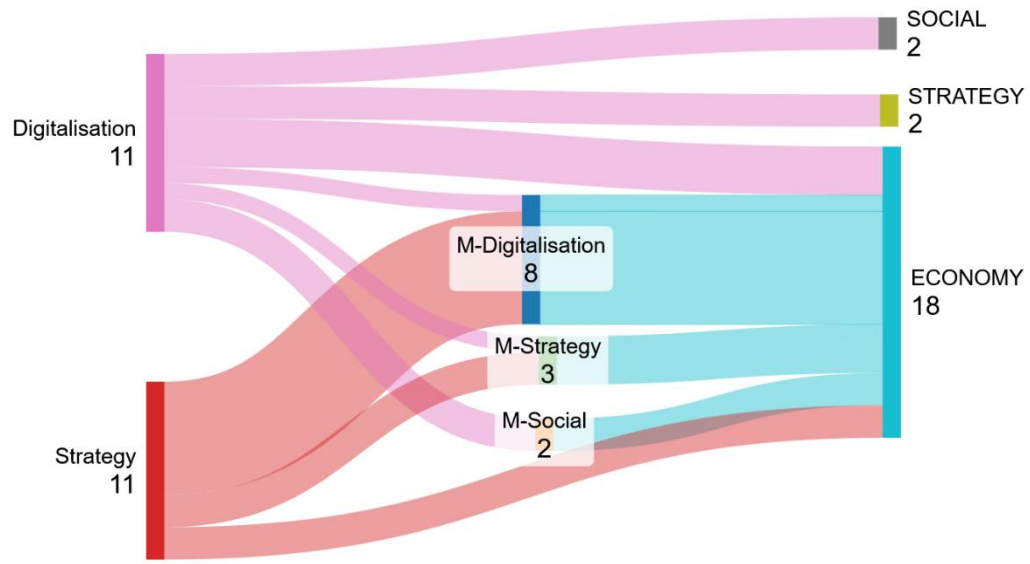
Qualitative studies complement the cluster by framing digitalisation on a meta-level and connecting it with sustainable development, business models, and systemic transformation (Acciarini et al., 2022; George et al., 2021). Performance is rarely central, appearing primarily in sustainability performance monitoring (Contini et al., 2023). The outcome perspective connects digitalisation to broader development goals (Guandalini, 2022; Santarius & Wagner, 2023), while at the operational level, digital technologies shape sustainable manufacturing (Agrawal, Majumdar, et al., 2023; Ching et al., 2022) and supply chain management (Rejeb & Rejeb, 2020; A. K. Sahu et al., 2023). Developing sustainable capabilities, defined as balancing short-term financial objectives with long-term transformation in production systems, is one key aspect of this transition (Kumar et al., 2021) as is innovation (Ghobakhloo et al., 2021).

(Digitally driven) Economic Efficiency

This cluster includes 22 variables from ten studies investigating how digital technologies improve decision-making, resource efficiency, and firm performance. Figure 7 visualises the contribution of digitalisation and strategy to profitability, operational

optimisation, and competitiveness, and their interaction with social and strategic sustainability dimensions through key social, strategic and digital mediating factors.

Figure 7: Cluster Impact Map – (Digitally driven) Economic Efficiency



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right (more detailed in the table below). "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Table 4: Cluster Profile – (Digitally driven) Economic Efficiency

Source	Independent Variable	Mediator [MODerator]	Dependent Variable	Code Pattern
Ghobakhloo, Asadi, et al., 2023	Intelligent Automation Implementation (+)		ECONOMIC SUSTAINABILITY PERFORMANCE	*DT →EM*
Al-Khatib, 2023	Industrial Internet of Things (IIoT) (+)	Supply Chain Visibility (+)	ECONOMIC PERFORMANCE	*DT → ST →EM*
Li, 2022	Digital Transformation (+)	[MOD: Market Turbulence (+)]		*DT →EM*
Li et al., 2022	Digital Technology Adoption (+)	Digital Technology Capability (+) [MOD: Digital Strategy (+)]		*DT → DT →EM*
Wang et al., 2023	Digital Orientation (+)	Supply Chain Internal Collaboration* (+)		*DT → SC* →EM*
		Supply Chain External Collaboration* (+)		*DT → SC* →EM*
			SUPPLY CHAIN INTERNAL COLLABORATION	*DT →SC**
			SUPPLY CHAIN EXTERNAL COLLABORATION	*DT →SC**
	AI-IoT Adoption		SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →EM*

Nayal et al., 2021[17]	**Following factors are supply chain flexibility factors:** Information Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Organisational Flexibility	AI-IoT Adoption (+)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Procurement Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Logistics Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Product Development Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Manufacturing Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
	Marketing Flexibility	AI-IoT Adoption (0)	SUPPLY CHAIN FIRM PERFORMANCE (SCFP)	*DT →DT →EM*
Kristoffersen et al., 2021	Business Analytics Capability (+)	Resource Orchestration Capability (+)	FIRM PERFORMANCE*	*DT →ST →EM*
		Circular Economy Implementation (+)		*DT →ST →EM*
Ukko et al., 2019	Managerial Capability (0)	[MOD: Sustainability Strategy (0)]	FINANCIAL PERFORMANCE	*ST →EM*
	Operational Capability (0)	[MOD: Sustainability Strategy (+)]		*ST →EM*
Ardito et al., 2021	Digital Orientation (+)		INNOVATION PERFORMANCE	*DT →ST*
Ardito, 2023	Cloud Computing (0)		SUSTAINABLE INNOVATION (SOCIAL AND ENVIRONMENTAL)	*DT →ST*

Qualitative Insights (not statistically tested)		
Source	**Impact Factor**	**Outcome Factor**
none		

Note: The direction and quality of effects is indicated as follows: (+) positive effect; (-) negative effect; (0) not significant.

Across the reviewed studies, a consistent narrative emerges: Digital transformation and technology adoption play pivotal roles in enhancing ECONOMIC (SUSTAINABILITY) PERFORMANCE, though the nature and magnitude of these effects vary depending on mediating factors, strategic alignment, and organisational capabilities. ECONOMIC (SUSTAINABILITY) PERFORMANCE stands out as the most comprehensive indicator of economic efficiency, reflecting both profitability and long-term sustainability. Ghobakhloo, Asadi, et al. (2023) show that intelligent automation implementation improves ECONOMIC SUSTAINABILITY PERFORMANCE, underscoring how automation can combine profitability with sustainability objectives. Similarly, Li (2022) demonstrates that digital transformation positively affects ECONOMIC PERFORMANCE, particularly under conditions of market turbulence, indicating that external pressures can amplify the

[17] Nayal et al. (2021) additionally examine the impact of each supply chain flexibility factor on AI-IoT adoption (coded relationship *ST → DT*). However, as these relationships fall outside the scope of our study, they have been excluded from the analysis.

economic benefits of digital initiatives. Extending this perspective, Li et al. (2022) reveal that digital technology adoption enhances ECONOMIC PERFORMANCE, with the relationship strengthened by digital technology capabilities and moderated by a firm's digital strategy. These findings emphasise the importance of both internal competencies and external conditions to maximising the economic impact of digitalisation.

The enhancement of SUPPLY CHAIN COLLABORATION – both internally and externally – is another central pathway through which digital technologies boost economic efficiency. Wang et al. (2023) find that digital orientation fosters INTERNAL AND EXTERNAL SUPPLY CHAIN COLLABORATION, which in turn improves ECONOMIC PERFORMANCE. Firms with a strong digital orientation adopt advanced technologies such as big data analytics and AI, enabling better communication, information sharing, and responsiveness. Other benefits include improved collaboration, streamlined operations, reduced costs, and enhanced market competitiveness, highlighting the interdependence between digital readiness and operational efficiency. At a more granular level and with the help of a complex structural equation modelling, Nayal et al. (2021) focus on SUPPLY CHAIN FIRM PERFORMANCE (SCFP), testing if supply chain flexibility factors in all areas of the firm improve SCFP when mediated by AI-IoT adoption. They find that most flexibility aspects influence the integration of AI and IoT, which in turn has a positive influence on SCFP. This finding illustrates how digital tools enable real-time data collection and adaptive decision-making, allowing firms to respond swiftly to market changes and improve financial results. However, there is no evidence for complex relationships. Kristoffersen et al. (2021) further highlight that business analytics capabilities do not directly influence FIRM PERFORMANCE but instead increase a firm's resource orchestration capability and the ability to succeed in the circular economy. This dual mediation pathway underscores the importance of data-driven resource management and sustainability-oriented practices in achieving economic gains.

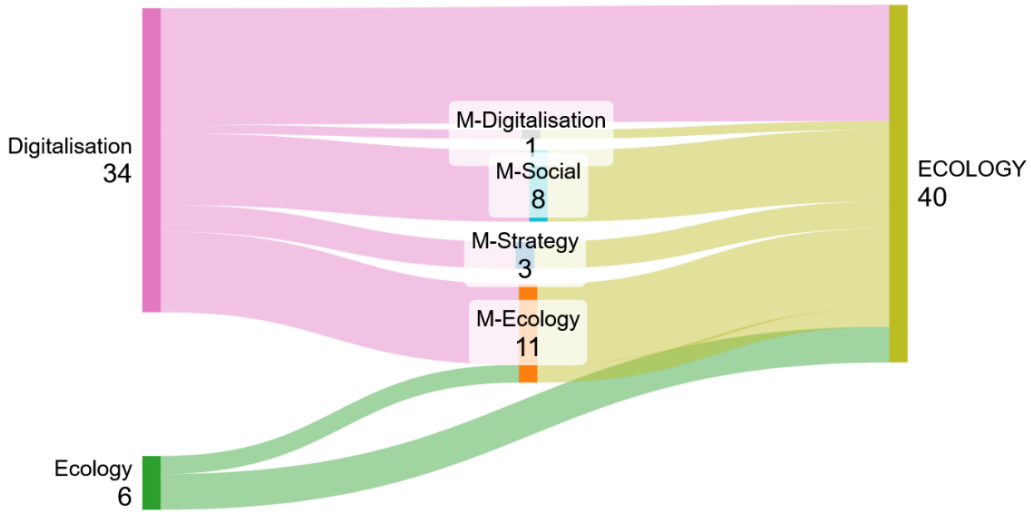
While most studies report positive impacts, some findings suggest that traditional organisational capabilities are insufficient to drive economic performance without digital integration. For instance, Ukko et al. (2019) find no significant direct relationship between managerial or operational capabilities and FINANCIAL PERFORMANCE, implying that conventional skills and processes should be augmented with digital competencies to yield economic benefits. This distinction reinforces the notion that digital transformation is not merely about adopting new technologies but about integrating them into broader

strategic and operational frameworks. INNOVATION PERFORMANCE represents another important dimension of economic efficiency. Ardito et al. (2021) demonstrate that digital orientation positively influences the innovation capabilities of firms, enabling the development of new products and processes that drive profitability. This relationship highlights the critical role of digital readiness in sustaining long-term economic competitiveness, with innovation serving as a key mechanism through which digitalisation translates into financial success. Another study of Ardito (2023) focuses on impact of concrete technology, namely cloud computing, on sustainable innovation, which is proven to be insignificant.

(Digitally driven) Ecologic Efficiency

Comprising 40 relationships that appear in ten studies, this cluster focuses on how digital technologies enable resource optimisation, emissions reductions, and improved environmental performance. Figure 8 visualises this strong, direct relationship, which is reinforced by mediating effects of social and strategic dimensions.

Figure 8: Cluster Impact Map – (Digitally driven) Ecologic Efficiency



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right (more detailed in the table below). "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Table 5: Cluster Profile – (Digitally driven) Ecologic Efficiency

Source	Independent Variable	Mediator [MODerator]	Dependent Variable	Code Pattern
Ferreira et al., 2023	Cloud Computing (+)		ENVIRONMENTAL SUSTAINABILITY	*DT →EL*
	Blockchain (+)			*DT →EL*
	Robotics (+)			*DT →EL*
	Big Data Analytics (+)			*DT →EL*
	AI (+)			*DT →EL*
Ghobakhloo, Asadi, et al., 2023	Intelligent Automation Implementation (+)		ENVIRONMENTAL SUSTAINABILITY PERFORMANCE	*DT →EL*
Li, 2022	Digital Transformation (+)	[MOD: Market Turbulence (+)]	ENVIRONMENTAL PERFORMANCE	*DT →EL*
Li et al., 2022	Digital Technology Adoption (+)	Digital Technology Capability (+) [MOD: Digital Strategy (+)]	ENVIRONMENTAL PERFORMANCE	*DT →DT →EL*
Al-Khatib, 2023	Industrial Internet of Things (IIoT) (+)	Supply Chain Visibility (+)		*DT →ST →EL*
Wang et al., 2023	Digital Orientation (+)	Supply Chain Internal Collaboration* (+)		*DT →SC* →EL*
		Supply Chain External Collaboration* (+)		*DT →SC* →EL*
Lerman et al., 2022	**All independent variables as part of the Smart Supply Chain:** Digital Transformation Strategy (+)		GREEN PERFORMANCE	*DT →EL*
	Base Digital Technologies (0)			*DT →EL*
	Front-end Technologies (0)			*DT →EL*
	Digital Transformation Strategy (+)	Customer Relationship (0)		*DT →SC →EL*
	Base Digital Technologies (0)			*DT →SC →EL*
	Front-end Technologies (+)			*DT →SC →EL*
	Digital Transformation Strategy (+)	Supplier Relationship (+)		*DT →SC →EL*
	Base Digital Technologies (0)			*DT →SC →EL*
	Front-end Technologies (+)			*DT →SC →EL*
	Digital Transformation Strategy (0)	Green Packaging (0)		*DT →EL →EL*
	Base Digital Technologies (0)			*DT →EL →EL*
	Front-end Technologies (0)			*DT →EL →EL*
	Digital Transformation Strategy (0)	Green Manufacturing (+)		*DT →EL →EL*
	Base Digital Technologies (0)			*DT →EL →EL*
	Front-end Technologies (0)			*DT →EL →EL*
	Digital Transformation Strategy (+)	Green Purchasing (0)		*DT →EL →EL*
	Base Digital Technologies (0)			*DT →EL →EL*
	Front-end Technologies (+)			*DT →EL →EL*

Wei et al., 2023	Supplier Low Carbon Collaboration (+)		CARBON PERFORMANCE	*EL→EL*
	Customer Low Carbon Collaboration (+)			*EL→EL*
	Eco-control systems (+)	Supplier Low-carbon Performance (+)		*EL→EL →EL*
		Customer Low-carbon Collaboration (+)		*EL→EL →EL*
	Eco-control Systems (+)	[MOD: Organizational Unlearning (0)]	SUPPLIER LOW-CARBON PERFORMANCE	*EL→EL*
	Eco-control Systems (+)	[MOD: Organizational Unlearning (+)]	CUSTOMER LOW-CARBON COLLABORATION	*EL→EL*
Di Maria et al., 2020	Smart-manufacturing Technologies (+)	Supply Chain Integration (SCI) (+)	CIRCULAR ECONOMY	*DT →ST →EL*
	Data-processing Technologies (0)			*DT →ST →EL*
	Smart-manufacturing Technologies (+)			*DT→EL*
	Data-processing Technologies (+)			*DT→EL*
Findik et al., 2023	Industry 4.0 Technologies (+)		CIRCULAR ECONOMY (PRACTICES)	*DT→EL*

Qualitative Insights (not statistically tested)		
Source	**Impact Factor**	**Outcome Factor**
Neligan et al., 2023	Digitalization	Efficiency of Circular Business Models
Rusch et al., 2022	Digital Technology	Sustainable Product Management in Circular Economy
Agrawal, Surendra Yadav, et al., 2023		Circularity in Supply Chains
Okorie et al., 2023		Net Zero Manufacturing Emissions
Böttcher et al., 2023		Ecological Sustainability in Business Models
He et al., 2023		Enterprise Green Strategy Evolution
Kristoffersen et al., 2020		Implementation of Circular Economy
Liu et al., 2022	Digital Functions	Circular Economy (Practices)
Neri et al., 2023	Digital-enabled Dynamic Capabilities	
Dwivedi & Paul, 2022	Digital Supply Chains	
Gupta & Singh, 2021	Industry 4.0 Implementation	
Ćwiklicki & Wojnarowska, 2020	Industry 4.0	
Lopes de Sousa Jabbour et al., 2018		
Sahu et al., 2022	Industry 4.0	Circular Economy
Parmentola et al., 2022	Blockchain	Environmentally Sustainable Development Goals
Yadav et al., 2023	Lean, Green and Digital Technologies	Net Zero Emissions

Note: The direction and quality of effects is indicated as follows: (+) positive effect; (-) negative effect; (0) not significant.

The reviewed studies consistently emphasise digitalisation as a key enabler of green efficiency, focussing on various dimensions of ecological outcomes: While ENVIRONMENTAL SUSTAINABILITY (Ferreira et al. 2023) reflects a company's overarching environmental objectives, ENVIRONMENTAL SUSTAINABILITY PERFORMANCE (Ghobakhloo, Asadi, et al., 2023a), and ENVIRONMENTAL PERFORMANCE (e.g., Li, 2022) focus on measurable environmental results. GREEN PERFORMANCE (Lerman et al., 2022) translates these overarching goals into concrete operational practices, particularly in supply chain and production contexts. Ferreira et al. (2023) examine how various digital technologies – including cloud computing, blockchain, robotics, big data analytics, and artificial intelligence – contribute to ENVIRONMENTAL SUSTAINABILITY, highlighting their positive effects on reducing emissions, optimising resource use, and improving environmental management systems. Building on this, Ghobakhloo et al. (2023) emphasise the importance of intelligent automation, which significantly enhances ENVIRONMENTAL SUSTAINABILITY PERFORMANCE by enabling precise environmental monitoring and real-time adjustments to operations. Li (2022) and Li et al. (2022) further explore how broader digital transformation efforts and digital technology adoption improve ENVIRONMENTAL PERFORMANCE, particularly under conditions of market turbulence. Their studies underscore the role of digital capabilities as both enablers and moderators of environmental outcomes, with digital strategies amplifying these positive effects. Connecting these broader environmental goals to operational execution, Lerman et al. (2022) find that advanced technologies like digital transformation strategies and front-end technologies positively influence GREEN PERFORMANCE by strengthening customer and supplier relationships, while foundational digital technologies have neutral effects. This underscores the need for the strategic integration of digital tools into organisational processes to achieve environmental goals. Similarly, Wang et al. (2023) highlight the importance of supply chain collaboration – both internal and external – in enhancing ENVIRONMENTAL PERFORMANCE, further bridging the gap between overarching sustainability goals and operational outcomes.

Carbon performance is another critical focus, particularly in studies examining supply chain collaborations. Wei et al. (2023) highlight that low-carbon collaborations between suppliers and customers, supported by eco-control systems, significantly improve CARBON PERFORMANCE. These collaborations facilitate the exchange of environmental information, enabling firms to implement low-carbon strategies more

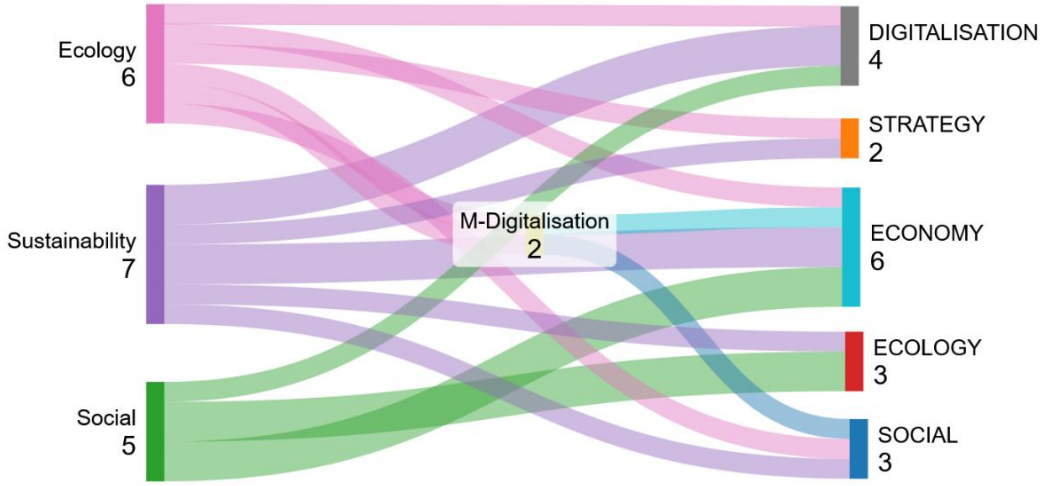
effectively. However, the relationship is moderated by organisational unlearning: while unlearning strengthens CUSTOMER LOW-CARBON COLLABORATION, it shows a non-significant effect on SUPPLIER COLLABORATION. This finding underscores the importance of organisational culture and adaptability in maximising the ecological benefits of digital technologies. The CIRCULAR ECONOMY (PRACTICES) represents another prominent outcome variable. Di Maria et al. (2022) find that smart-manufacturing technologies enhance circular economy outcomes, particularly when mediated by supply chain integration (SCI). However, data-processing technologies show mixed results, with direct effects on circular economy practices but no significant impact when SCI is considered as a mediator. Findik et al. (2023) further support the positive relationship between Industry 4.0 technologies and circular economy practices, emphasising the role of advanced manufacturing and digital integration in promoting resource efficiency.

The qualitative contributions in this cluster emphasise the role of digital technologies in advancing circular economy (practices) and improving ecological sustainability, especially net-zero emissions. Different scholars (Ćwiklicki & Wojnarowska, 2020; Gupta & Singh, 2021; Lopes de Sousa Jabbour et al., 2018) examine how Industry 4.0 applications promote resource efficiency and circular economy adoption. Liu et al. (2022) and Neri et al. (2023) highlight digital functions and dynamic capabilities as enablers of circular business models. Okorie et al. (2023) and Yadav et al. (2023) address how digital technologies contribute to net-zero manufacturing and emissions reduction, while Parmentola et al. (2022) explores blockchain's potential for achieving environmentally sustainable development goals. The strategic integration of these technologies is viewed as key to achieving ecological outcomes (Böttcher et al., 2023; Neligan et al., 2023).

Sustainability-Driven Business and Digitalisation

Encompassing 18 relationships in eight studies, this cluster explores how corporate sustainability initiatives align with digital technologies and with the integration of sustainability into core business strategies. Figure 9 visualises how ecological, social, and sustainability-related factors interact with economic and strategic dimensions. Digitalisation emerges as both an outcome and a mediator.

Figure 9: Cluster Impact Map – Sustainability-Driven Business and Digitalisation



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right (more detailed in the table below). "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Table 6: Cluster Profile – Sustainability-Driven Business and Digitalisation

Source	Independent Variable	Mediator	Dependent Variable	Code Pattern
Xu et al., 2023	Sustainable Performance (+)		ECONOMIC PERFORMANCE	*SU → EM*
Wang et al., 2023	Supply Chain Internal Collaboration* (+)			*SC* → EM*
	Supply Chain External Collaboration* (+)			*SC* → EM*
Torrent-Sellens et al., 2023	Environmental Assets (+)			*EL → EM*
	Environmental Assets (+)	Industry 4.0 (+)		*EL → DT →EM*
Broccardo et al., 2023	Sustainability Performance (+)		COMPANY PROFITABILITY PERFORMANCE	*SU → EM*
Pînzaru et al., 2022	Sustainability Practices (+)		BENEFITS PERCEIVED BY COMPANIES	*SU → ST*
Ardito et al., 2021	Environmental Orientation (+)		INNOVATION PERFORMANCE	*EL→ST*
Torrent-Sellens et al., 2023	Environmental Assets (+)		SOCIAL PERFORMANCE	*EL → SC*
	Environmental Assets (+)	Industry 4.0 (+)		*EL → DT →SC*
Xu et al., 2023	Sustainable Performance (+)			*SU → SC*
	Sustainable Performance (+)		ENVIRONMENTAL PERFORMANCE	*SU → EL*
Wang et al., 2023	Supply Chain Internal Collaboration* (+)			*SC* → EL*
	Supply Chain External Collaboration* (+)			*SC* → EL*
Ghobakhloo, Asadi, et al., 2023	**Environmental factors*:** Environmental Turbulence (0) External Stakeholder Pressure (0)		INTELLIGENT AUTOMATION IMPLEMENTATION	*EL* → DT*
	Human factors*: Social Capital Competency (+) Employee Socio-Behavioral Concerns (-) Management Digitalization Competency (+)			*SC* → DT*

Pînzaru et al., 2022	Internal Sustainability Factors (+)		DIGITALIZATION	*SU → DT*
	External Sustainability Factors (+)			*SU → DT*
Qualitative Insights (not statistically tested)				
Source	**Impact Factor**	**Outcome Factor**		
Feroz et al., 2023	Dynamic Capabilities	Sustainable Digital Transformation		
Pauliuk et al., 2022	Co-design of Digital Transformation and Ecological Sustainability			
Isensee et al., 2020	Organizational Culture	Environmental Sustainability		
Mukhuty et al., 2022	Social Responsibility and HR Practices	Industry 4.0 Development		
Benešová et al., 2021	Green Strategies	Maturity Models for Industry 4.0		
Ribeiro et al., 2021	Implementing (Digital) Social Innovation in Developing a Tool to Support Product Strategy.			
Demir et al., 2023	Smartness and Sustainability Aspects	Supply Chain Operations		
zu Knyphausen-Aufseß & Santarius, 2021	Role of Firms	Digitalization		
Niehoff, 2022	Sustainability Reports			
Goede, 2021	Sustainability	Data-driven Decision Making in Organisations and Business Intelligence Systems		
Zarte et al., 2022	Sustainability Aspects	Knowledge Framework for the Collection of Data		
Sætra, 2023	AI ESG Protocol			

Note: The direction and quality of effects is indicated as follows: (+) positive effect; (-) negative effect; (0) not significant.

Across the cluster, ECONOMIC PERFORMANCE emerges as a central outcome variable. Studies show that sustainable performance positively influences economic outcomes (Xu et al., 2023), while sustainability practices improve BENEFITS PERCEIVED BY COMPANIES (Pînzaru et al., 2022). Similarly, environmental assets are linked to better economic performance, with Industry 4.0 partially mediating this relationship, underscoring the role of advanced technologies in translating sustainability efforts into profitability (Torrent-Sellens et al., 2023). Broccardo et al. (2023) further demonstrate that improved sustainability performance is associated with increased COMPANY PROFITABILITY PERFORMANCE, emphasising the economic value of sustainability-oriented strategies. In terms of INNOVATION PERFORMANCE, Ardito et al. (2021) find that an environmental orientation enhances a firm's capacity to innovate, particularly when aligning sustainability objectives with product and process development. SOCIAL PERFORMANCE is similarly impacted by sustainability drivers. Xu et al. (2023) report that sustainable performance contributes to enhanced SOCIAL PERFORMANCE, while Torrent-Sellens et al. (2022) show that environmental assets improve social outcomes both directly and through the partial mediation of Industry 4.0. ENVIRONMENTAL PERFORMANCE and GREEN PERFORMANCE also feature prominently. Xu et al. (2023) demonstrate that sustainable

performance drives environmental improvements, whereas Wang et al. (2023) highlight the role of supply chain internal and external collaboration in achieving better environmental outcomes. Lerman et al. (2022) find no effect of customer relationships on GREEN PERFORMANCE, as long as they are strategically integrated into sustainability initiatives.

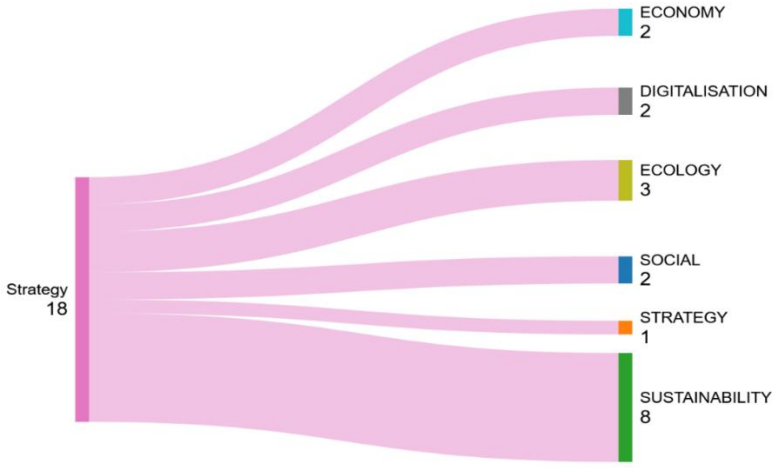
A closer look into Sustainability-Driven Digitalisation reveals how sustainability considerations influence technological adoption. Pînzaru et al. (2022) show that internal and external sustainability factors drive DIGITALISATION efforts, emphasising the role of organisational culture and regulatory pressures. Ghobakhloo, Asadi, et al. (2023) further illustrate how human factors such as social capital competency, management digitalisation competency, and employee concerns shape the extent of INTELLIGENT AUTOMATION IMPLEMENTATION, whilst external stakeholder pressure and environmental turbulence seem to have no effect.

The qualitative contributions in this cluster explore how sustainability considerations shape digital transformation and organisational strategies. Feroz et al. (2023) conceptualise digital transformation as more than technological adoption; instead, they argue that, to achieve long-term sustainability, it requires fundamental shifts in processes, structures, and mindsets, with leadership commitment and strategic alignment serving as key enablers. Ribeiro et al. (2024) explore Digital Social Innovation, identifying key enablers such as social value creation, stakeholder involvement, digital ecosystems, economic sustainability, and risk management. Benešová et al. (2021) examine how green strategies influence Industry 4.0 maturity models. Organisational culture and stakeholder engagement emerge as central themes, with Isensee et al. (2020) and Mukhuty et al. (2022) addressing how internal drivers and HR practices support sustainable digital transformation. Niehoff (2022) examines sustainability reporting as a tool for integrating digitalisation within corporate sustainability strategies and highlights how data-driven decision-making in organisations depends on factors such as clear vision, management support, and high-quality user access tools, which influence the effectiveness of sustainable business intelligence systems. Demir et al. (2023) focus on smartness and sustainability aspects in supply chain operations, and Sætra (2023) introduces an AI ESG protocol to assess sustainability impacts.

Strategy as a Catalyst for Sustainable Transformation

This cluster comprises 18 relationships that appear in six quantitative empirical studies highlighting the pivotal role of sustainability-oriented strategies and practices in shaping firm performance, process improvements, and technological adoption. Figure 10 illustrates the strong link between strategy and the ecological, economic, social, and digitalisation dimensions, emphasising how the integrative function of strategy aligns diverse sustainability efforts emphasising its integrative function in aligning diverse sustainability efforts at the corporate level.

Figure 10: Cluster Impact Map – Strategy as a Catalyst for Sustainable Transformation



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right (more detailed in the table below). "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Table 7: Cluster Profile – Strategy as a Catalyst for Sustainable Transformation

Source	Independent Variable (positive influence)	Mediator	Dependent Variable	Code Pattern
Chatzistamoulou, 2023	**Firm specific heterogeneity:** Decreased Turnover (-) Firm Size (-) Firm Age (0)		SUSTAINABILITY TRANSITION	*ST → SU*
	Business operational problems: Corruption (0) Complexity of Administrative Procedures (0) Fast Changing Legislation (0) Inadequate Infrastructure (0) Tax Rates (0) Access to Financing (0)			*ST → SU*
	The country attitude towards sustainability: Resource Productivity (-) Renewable Energy Use (-) Eco Innovation Index (+)			*ST → SU*

	Environmental Protection Expenditure at the National Level (+)			
	Attitude towards business environment: Digital Competitiveness Ranking (+) Regulation (0) Corruption Perception Index (-)			*ST → SU*
	Public procurement processes: Perceived Corruption in Public Procurement Nationally (+) Public Procurement Participation (+)			*ST → SU*
Tasleem et al., 2019	Total Quality Management (+)		CORPORATE SUSTAINABILITY PERFORMANCE	*ST → SU*
	Technology Management (0)			*ST → SU*
	Total Quality Management (+)		ENVIRONMENTAL SUSTAINABILITY PERFORMANCE	*ST → EL*
	Technology Management (0)			*ST → EL*
	Total Quality Management (+)		SOCIAL SUSTAINABILITY PERFORMANCE	*ST → SC*
	Technology Management (0)			*ST → SC*
	Total Quality Management (+)		ECONOMIC SUSTAINABILITY PERFORMANCE	*ST → EM*
	Technology Management (+)			*ST → EM*
Broccardo et al., 2023	Company Size (+)		SUSTAINABILITY IMPLEMENTATION AND RELATED COMPANY PERFORMANCE	*ST → SU*
Ardito et al., 2021	Interaction between Digital Orientation and Environmental Orientation (0)		INNOVATION PERFORMANCE	*ST → ST*
Di Maria et al., 2020	Supply Chain Integration (+)		CIRCULAR ECONOMY	*ST → EL*
Ghobakhloo, Asadi, et al., 2023	**Technological Factors:** Investment Risk (0) Cybersecurity Risk (0) Integrability (0) Strategic Value (+)		INTELLIGENT AUTOMATION IMPLEMENTATION	*ST → DT*
	Organisational factors: Absorptive Capacity (+) Digitalization Technical Competency (+) Resource Availability (+)			*ST → DT*

Qualitative Insights (not statistically tested)		
Source	**Impact Factor**	**Outcome Factor**
Allal-Chérif et al., 2023	Disruptive Strategy	Successful Sustainable Entrepreneurship
	Social Product Innovation	

	Digitalization	
Dwivedi et al., 2023	Digital Supply Chains	Circular Economy Adoption
Grunwald, 2022	Sustainability Co-Creation	Digitalised Global Value Chains
Haftor & Climent, 2021	Innovative Offerings in Industry	Environmental Sustainability
Hristov & Appolloni, 2021	Sustainable Development	Integration of Stakeholder Perception in Performance Management System
	Organisational Drivers	
	Digital Transformation	
	Cultural Context	

Note: The direction and quality of effects is indicated as follows: (+) positive effect; (-) negative effect; (0) not significant.

The concept of SUSTAINABILITY TRANSITION emerges prominently in Chatzistamoulou (2023), who investigates how firm-specific factors (e.g., decreased turnover, smaller size, older firm age) and business operational challenges (such as corruption, administrative complexity, and access to financing) hinder the adoption of sustainability-focused strategies. Interestingly, external conditions – such as a country's attitude toward sustainability, renewable energy use, and eco-innovation – exert mixed influences, with positive effects obtained from eco-innovation indices but negative or neutral impacts from resource productivity and corruption perceptions. Public procurement participation positively contributes to the transition, underscoring how external institutional environments shape firms' sustainability pathways. Closely tied to this theme, CORPORATE SUSTAINABILITY PERFORMANCE is examined by Tasleem et al. (2019), who find that Total Quality Management (TQM) significantly improves economic, environmental, and social sustainability outcomes, whereas Technology Management shows no direct effect, unless it is mediated by TQM. These findings emphasise the importance of structured management approaches in enhancing sustainability performance across all dimensions. Furthering this, Broccardo et al. (2023) analyse SUSTAINABILITY IMPLEMENTATION AND RELATED COMPANY PERFORMANCE, showing that larger company size positively influences the integration of sustainability into business operations. Their findings highlight how organisational capacity and resource availability facilitate the embedding of sustainability practices, leading to improved performance outcomes.

Innovation-driven perspectives also feature in this cluster. Ardito et al. (2021) explore INNOVATION PERFORMANCE and reveal that while environmental orientation positively affects innovation outcomes, the interaction between digital and environmental orientation yields mixed effects – partially supporting the notion that pursuing both

simultaneously can strain organisational resources, potentially dampening innovation gains. Likewise, SUSTAINABLE INNOVATION – particularly when it addresses both social and environmental dimensions – is explored in studies emphasising the importance of strategic sustainability efforts in fostering innovative solutions. These innovations are shown to stem from firms adopting comprehensive sustainability strategies that align with broader environmental and social goals. The cluster also examines resource-related outcomes, particularly through the lens of circular economy initiatives. Di Maria et al. (2022) demonstrate that supply chain integration plays a crucial role in enhancing CIRCULAR ECONOMY practices, with smart manufacturing technologies positively influencing integration and, consequently, circular outcomes. However, the effect of data-processing technologies remains inconclusive, highlighting the varying effectiveness of digital solutions in sustainability contexts.

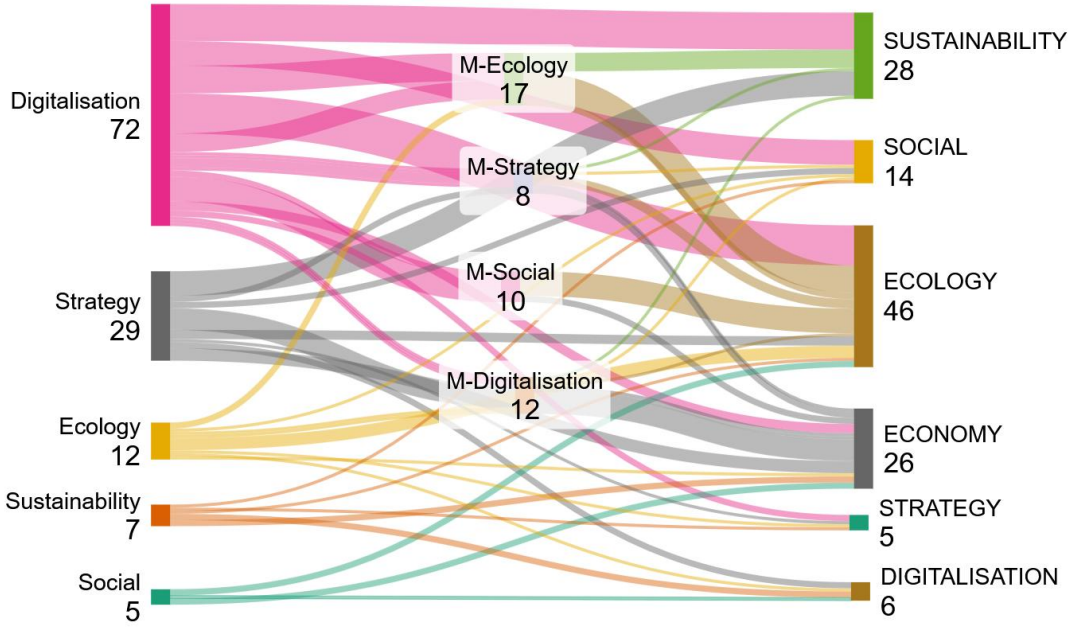
Finally, INTELLIGENT AUTOMATION IMPLEMENTATION emerges as a critical technological pathway influenced by sustainability drivers. Ghobakhloo, Asadi, et al. (2023) show that organisational factors (such as absorptive capacity and digitalisation competency) and technological considerations (like strategic value) positively affect automation adoption, enabling firms to better meet sustainability objectives through process efficiencies. Conversely, environmental factors (e.g., environmental turbulence) do not significantly influence adoption decisions, underscoring the predominance of internal capabilities over external environmental pressures in driving technological integration.

Qualitative studies underscore strategy as a key enabler for aligning digitalisation with sustainability goals. Dwivedi et al. (2023) explore digital supply chains as catalysts for circular economy adoption, while Allal-Chérif et al. (2023) highlight how disruptive strategies and social product innovation foster sustainable entrepreneurship. Innovative offerings from industrial organisations (Haftor & Climent, 2021) and sustainability co-creation within digitalised global value chains (Grunwald, 2022) emerge as pivotal themes. Organisational drivers and stakeholder integration are crucial for fostering adaptability and embedding sustainability into performance management systems (Hristov & Appolloni, 2022).

2.5 Discussion of Key Findings and Research Gaps

Guided by our research question we provide a structured overview of the research field, outlining key thematic and methodological patterns. The sample impact map (Figure 11) illustrates an interconnected yet fragmented research landscape, where digitalisation frequently serves as a catalyst for corporate sustainability outcomes. The map reflects the field's strong focus on technology-driven efficiency and environmental improvements. Strategy also plays a pivotal role, acting as a bridge between digitalisation and economic and sustainability outcomes. In contrast, connections to social dimensions remain comparatively limited, indicating an underexplored area in current research. Building on this cross-topic perspective, the following discussion systematically explores key findings and research gaps derived from the cluster analysis, and offers insights into prevailing research foci, thematic imbalances, and potential future research directions.

Figure 11: Aggregated Cluster Impact Map



Note: The figure depicts all the relationships examined in this cluster, regardless of the direction and significance of the effect. Categories of independent variables are on the left; categories of dependent variables are on the right. "M" denotes mediators; moderators are not depicted. Numbers represent the frequency of variables in each category. The thickness of the connecting lines indicates the frequency of a given relationship.

Fragmented Sustainability Perspectives

While all the studies in our sample inherently address environmental, social, economic, and strategic dimensions, due to the breadth of our search string, only a few explicitly adopt a multi-dimensional sustainability approach. Instead, sustainability

dimensions are predominantly analysed individually or in specific combinations, such as the alignment of strategy with ecology or the economic impacts of circular economy practices. This fragmented approach limits a comprehensive understanding of how sustainability dimensions interact. Though necessary for analytical clarity, such separation underrepresents systemic interdependencies, potentially overlooking rebound effects and synergies between sustainability goals. Future research should develop integrated frameworks that incorporate ecological, economic, and social sustainability dimensions, and explore multi-level models that capture sustainability interdependencies in corporate settings to overcome the conceptual divide between research on CS outcomes, and on CS approaches and knowledge (Montiel et al., 2020; Pranugrahaning et al., 2021). As research increasingly investigates the intersection of CS and digitalisation, it has become all the more important to account for interdependencies between the two, and to avoid overlooking rebound effects and synergies. Additionally, trans- and interdisciplinary studies could provide valuable insights into how trade-offs between digitalisation, strategy, and sustainability are managed in practice.

Digitalisation as Main Enabler of Sustainability

Consistent with prior literature positioning digitalisation as a transformative force (e.g., Flyverbom et al., 2019), our findings show that digitalisation is primarily examined as an enabler rather than as a sustainability goal. Studies highlight how the AI, IoT, big data analytics, and digital business transformation contribute to achieving sustainability objectives. Particularly when the sustainability objectives are efficiency-driven. This aligns with the concept of digital sustainability, which describes the synergy effects of digitalisation on sustainability goals (George et al., 2021; Guandalini, 2022). Yet, research remains largely efficiency-oriented, focusing on operational gains rather than systemic sustainability transformations. There is a notable gap in understanding the organisational and contextual factors – such as culture, governance, and strategic alignment – that influence digital sustainability outcomes. Longitudinal studies assessing the lasting impacts of digitalisation on sustainability, and sector-specific analyses in underrepresented industries, could contribute valuable knowledge. Moreover, the concept of sustainable digitalisation – embedding sustainability principles in digital infrastructure development – remains largely unexplored in the sample, highlighting the need for research on how to prevent unintended social and environmental consequences through appropriate governance.

Strategy as a Key Factor in Sustainability Integration

Anchoring sustainability in corporate strategy is well-established in CS research (Amini & Bienstock, 2014) and therefore reflected in our findings. Strategy emerges as a primary driver and frequently as a mediator between digitalisation and sustainability, underscoring its function as organisational infrastructure that translates technological capabilities into sustainability outcomes. Mechanisms such as resource orchestration, supply chain visibility, and technology management illustrate how strategic frameworks enable alignment between digital tools and sustainability goals. This finding mirrors the distinction between outcome-driven and implementation-driven sustainability research (Van Zanten & Van Tulder, 2021). However, it also reveals a blatant gap: while many studies focus on sustainability performance outcomes, only a few explore how sustainability is institutionalised within strategic planning processes. Future research should not only investigate how strategic capabilities shape long-term sustainability performance, but also examine how governance structures support sustainability-oriented decision-making, and how strategic planning itself can be positioned as a sustainability objective. Particular attention should be paid to the integration of social and governance dimensions alongside ecological and economic priorities.

Dominance of Ecology and Efficiency-related Topics

Our review reveals the dominance of ecology-related topics, represented by the largest thematic area of both mediator and outcome variables. This prominence suggests that sustainability research, particularly in connection with digitalisation, is largely shaped by efficiency-driven perspectives, focusing on measurable outcomes such as operational efficiency, resource optimisation, and ecological performance. Digitalisation is primarily examined in relation to these quantifiable aspects, aligning closely with regulatory frameworks that prioritise environmental compliance and cost-efficiency measures. Studies emphasise how digitalisation enables resource efficiency, closed-loop production, and sustainable supply chain operations, reinforcing the prevalence of technical and economic sustainability. While valuable, this focus often overlooks broader systemic transformations. Notably, the integration of digitalisation and sustainability remains prevalent in manufacturing contexts, where digital solutions have immediate operational impacts. Social and governance dimensions remain underexplored, even though regulations like the EU's Corporate Sustainability Reporting Directive (CSRD) suggests their growing importance, reshaping forthcoming corporate priorities. To

address these gaps, future research should move beyond efficiency-driven perspectives, exploring the mitigation of rebound effects through governance mechanisms and sustainability-oriented digital strategies (Bohnsack et al., 2022). Regulatory shifts like the CSRD offer opportunities to examine how ESG compliance affects long-term corporate investment decisions beyond regulatory obligations.

Shifting Research Priorities in Topics and Methods

The research landscape shows first signs of evolving beyond efficiency-driven, ecology-focused orientation. While themes like Industry 4.0, circular economy, and supply chain management remain prevalent, there is a noticeable tendency toward exploring strategic, social, and governance-related sustainability dimensions. The rising interest in sustainable digitalisation and its strategic implications suggests a broadening of research priorities. Until recently, sustainability research has been shaped by measurable, performance-oriented outcomes, often aligning with regulatory frameworks that emphasise environmental compliance and financial performance. This explains why ecological goals have received the most attention. Yet, regulatory developments like the CSRD are likely to elevate the prominence of strategy and social sustainability considerations (Tettamanzi et al., 2022). From a methodological perspective, the research field remains largely quantitative, frequently employing performance metrics to assess sustainability outcomes. This favours ecological and economic research areas where sustainability impacts can be quantified and directly linked to technological innovations. While such approaches offer clarity, they often fail to capture the complexity of sustainability transitions, particularly in social and governance contexts. Recent trends present an increasing methodological diversity by adopting mixed-methods and qualitative, especially case-based, approaches (e.g., Allal-Chérif et al., 2023; Feroz et al., 2023). Despite this diversification, systematic literature reviews continue to be a dominant approach in this field. Refining the focus and methodological positioning of such reviews could help guide future research more effectively towards specific directions (e.g., Krlev et al., 2025). To fully grasp the synergies and challenges in the context of digitalisation and sustainability, future research should embrace longitudinal studies, case-based research, and mixed-methods designs to provide a more comprehensive understanding of sustainability transitions and its dynamics. This expansion should not replace the ecological focus but complement it by integrating new perspectives alongside efficiency concerns.

2.6 Conclusion

The reviewed, and a wide array of identified drivers and barriers. The identified clusters reflect both the potential and the complexity of integrating digitalisation into corporate sustainability. This review highlights the dual role of digital technologies and strategic planning as enablers of systemic change while revealing gaps in understanding how these dimensions interact across different contexts. Digitalisation is often portrayed as an enabler of CS. The relative dominance of ecological topics further emphasises measurable, efficiency-driven aspects of sustainability, often linked to digital technologies and regulatory incentives. In contrast, social dimensions remain articles exhibit considerable variation in scope, methodological approaches, and perspectives. These divergences stem from differing conceptions of sustainability and digitalisation underexplored, with limited integration into sustainability frameworks. This synthesis underscores the need for an integrated approach that captures the interdependence between sustainability dimensions. Advancing this integration enables addressing the dynamic interplay of digitalisation, strategy, and sustainability. By building on these insights, future research can provide more comprehensive pathways for operationalising sustainability in a rapidly evolving digital landscape.

Table 8: Key Findings, Research Gaps and Future Directions

Key Findings	Research Gaps	Research Paths	Examples of Research Questions
Fragmented Sustainability Perspectives	Lack of integrated, multi-dimensional sustainability approaches	Develop integrated frameworks accounting for ecological, economic, and social sustainability dimensions	
		Explore multi-level models capturing sustainability interdependencies	
		Increase trans- and interdisciplinary studies	
	Predominant focus on isolated sustainability dimensions	Integrate sustainability perspectives across research disciplines	What strategies can promote a cross-disciplinary approach to sustainability research?
		Develop methodologies that address sustainability holistically	How can sustainability research frameworks be adapted for multi-dimensional analysis?
Dominance of Ecology and Efficiency-related Topics	Underrepresentation of social and governance dimensions	Expand empirical research beyond efficiency-driven approaches	How can social and governance sustainability dimensions be better integrated into corporate sustainability strategies?
		Investigate the impact of regulatory frameworks (e.g., CSRD) on corporate sustainability priorities	What are the long-term sustainability effects of efficiency-driven digitalisation?
	Overemphasis on efficiency-driven sustainability approaches	Develop interdisciplinary methodologies integrating ecological, economic, and social sustainability dimensions	How do different sustainability perspectives impact digital transformation strategies?
Digitalisation as Main Enabler of Sustainability	Limited research on digitalisation as a sustainability goal rather than an enabler	Investigate sector-specific applications of digital sustainability	What factors determine whether digitalisation leads to sustainability outcomes beyond efficiency improvements?
		Conduct longitudinal studies on digital sustainability transitions	What are the long-term sustainability implications of digital transformation?
	Insufficient focus on contextual factors affecting digital sustainability adoption	Explore governance mechanisms preventing unintended sustainability consequences	How do governance structures shape sustainable digital transformation strategies?

Strategy as a Key Factor in Sustainability Integration	Lack of research on institutionalising sustainability within corporate strategy	Explore strategy as a sustainability objective in itself	
Shifting Research Priorities in Topics and Methods	Predominance of quantitative, efficiency-focused models	Expand use of mixed-methods and qualitative research in sustainability studies	
		Bridge research design gaps in digital sustainability and social governance factors	How do evolving regulations shape corporate sustainability priorities and strategies?
	Limited use of mixed-methods and qualitative approaches	Investigate evolving corporate sustainability strategies amid increasing regulatory pressures	

CHAPTER 3 | Navigating Responsibility in the Digital Age: A Systematic Literature Review Comparing Corporate Digital Responsibility and Corporate Social Responsibility

Eva A. Jakob Sabrina Plaß

The impact of digital transformation on businesses and society has been researched in many disciplines and from multiple perspectives. Two prominent but fragmented research streams have started to explore how corporations can conduct digitalisation responsibly: Corporate Digital Responsibility (CDR) and Corporate Social Responsibility (CSR). However, the relationship between these two streams, and whether they converge or diverge, has been inconclusive. This paper conducts a systematic literature review by comparing and synthesising 55 articles on CDR, and 75 on the intersection of CSR and digitalisation. Our critical analysis discusses and highlights the differences in methodological approaches, the choice of topics, and conceptualisations adopted in these different research communities. This review contributes a more comprehensive understanding of CDR and CSR in an era of rapid digital change, and offers guidance and avenues for future research for scholars and practitioners seeking to navigate this evolving field.

Keywords: *Corporate Social Responsibility, Corporate Digital Responsibility, Digitalisation, Systematic Literature Review, Stakeholders*

JEL Codes: M10, M14, M15, O33

3.1 Introduction

Since the Facebook-Cambridge-Analytica scandal in 2018, ethical concerns surrounding corporate data privacy have gained increasing attention. Algorithmic decision-making remains opaque, while digital transformation raises social and environmental questions. Policymakers, companies, and scholars started discussing the role of companies in fostering socially responsible digitalisation (Flyverbom et al., 2019; Loebbecke & Picot, 2015). In response, Corporate Digital Responsibility (CDR) has emerged as a concept guiding ethical corporate behaviour in the digital domain. Initially promoted by consulting agencies (e.g., Accenture, Deloitte), multinational companies (e.g., Miele, Otto Group), and joint initiatives in Europe (Cooper et al., 2015; Kunicke, 2018), CDR has gained traction in academic discourse (Carl & Hinz, 2024; Knopf & Pick, 2023; Lobschat et al., 2021b). Given the disruptions, opportunities, and (unintended) consequences of digitalisation[18], scholars argue for the adoption of norms, especially with a view to data protection, IT security, algorithmic fairness, and the ethical use of technology synthesised in the concept of CDR (e.g., Lobschat et al., 2021; Mihale-Wilson et al., 2022). A commonly used definition positions CDR 'as the set of shared values and norms guiding an organization's operations with respect to the creation and operation of digital technology and data' (Lobschat et al., 2021:876).

Alongside the emerging debate about Corporate Digital Responsibility (CDR), which seeks to address the key question of companies' responsibility in relation to digitalisation, there are also parallel debates in research about the intersection of digitalisation and Corporate Social Responsibility (CSR) (e.g., Etter et al., 2019; Ghobakhloo et al., 2023). CSR broadly refers to a voluntary corporate commitment to society whereby companies proactively consider the economic, social, and environmental factors of their actions (Aguinis, 2011; Dahlsrud, 2008). Despite the growing body of research on both CDR and CSR in digital contexts, there have been no attempts to date to systematically compare these fields, or to raise potential concerns about conceptual redundancy arising from a fragmented discourse. Existing literature reviews primarily focus on definitions of CDR (Bednárová & Serpeninova, 2023) by mapping its contents and motivations (Knopf & Pick, 2023; Weber-Lewerenz & Traverso, 2024; S. S. K. Yadav

[18] While different terms like datafication, digitalisation and digital transformation vary across disciplines (Brenner & Hartl, 2021; Loebbecke & Picot, 2015), we refer to digitalisation as an umbrella term to account for the way organisations and stakeholders utilise digitised data and digital technologies, including the development of AI (Legner et al., 2017b).

& Mishra, 2022) and situating it either within specific research communities (Carl, 2023; Carl & Hinz, 2024) or within specific industries, such as Fintech and construction (Aldboush & Ferdous, 2023; Weber-Lewerenz & Traverso, 2024, for an overview of the reviews, see Appendix A1, p. 209). Covucci et al. (2024) review CDR research in relation to digital sustainability, while Atanasov et al. (2023) analyse research on the integration of digital technologies in CSR. These reviews underline the growing interest in and importance of CDR, and the need to integrate its fragmented research. In this context we argue that it is crucial for the systematic advancement of knowledge to delineate the boundaries between CDR and CSR in digital contexts, and to identify both their distinct contributions and their areas of convergence. Hence, we pose the following research question: *What are the essential differences between research on CDR and research that lies at the intersection of CSR and digitalisation?*

To enable us to systematically compare these two streams of literature, and to identify commonalities and differences in their methods, communities, and topics, we conduct a systematic literature review on, respectively CDR (n=55), and on the intersection of CSR and digitalisation (n=75). By using thematic coding, we identified three overarching topics in each of the streams. Based on the results or our comparative analysis, we reflect on CDR's relationship with CSR and critically discuss future research opportunities.

This study makes three key theoretical contributions to the growing body of research on CDR and CSR in the digital age. First, we offer a comparative analysis that clarifies the distinct contributions of CDR and CSR in the context of digitalisation. In terms of divergences, our findings reveal that CDR research actively engages with ethical challenges, such as AI integration, data security governance, and digital trust, whereas the main foci of the CSR literature are empirical concerns, such as performance measurement, the interaction between sustainability and digitalisation, or the role of CSR in digital communication. On the other hand, the two fields converge on some important themes, such as stakeholder engagement and customer focus. However, this can lead to conceptual overlaps and potential confusion. By systematically identifying these overlaps, this study helps delineate the boundaries between CDR and CSR, offering a clearer understanding of their respective roles in responsible digitalisation.

Second, the study advocates a closer integration of the two fields, urging scholars to reconcile the conceptual and methodological distinctions between CDR and CSR. While CDR's emerging focus on the digital domain is important, it should not ignore the foundational contributions of CSR, which already addresses the social, economic, and environmental implications of digitalisation (Flyverbom et al., 2019; González-Ramírez et al., 2024; G. Huang & Shen, 2024). The study stresses the importance of fostering dialogue between these two research streams to avoid a fragmented discourse and to promote a holistic approach to corporate responsibility in the digital age.

Third, the study highlights the need for more actionable, measurable frameworks in both CDR and CSR. While CDR research has largely been theoretical in nature, CSR research has been more practice-oriented, focusing on empirical issues like digital communication, and an integrated approach to sustainability (Bednárová & Serpeninova, 2023; Weber-Lewerenz & Traverso, 2024). Our analysis suggests the need to further develop validated metrics, and for future research to bridge the gap between theoretical discussions and tangible, responsible corporate practices in digitalisation. This would help companies to navigate more effectively the challenges of responsible digital transformation, while ensuring long-term holistic sustainability including social, environmental and economic dimensions.

3.2 Systematic Literature Review on CDR and the Intersection of CSR and Digitalisation

We conducted a systematic literature review of Corporate Digital Responsibility research, and of the research lying at the intersection of CSR and digitalisation. Following the approach by Denyer and Tranfield (2009), we applied the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) guidelines, which involved searching for relevant articles, selecting and evaluating sources, analysing and synthesising findings (see search protocol in Figures 2 and 3).

To identify relevant articles, we conducted two search procedures. First, as we were particularly interested in the emergence of the concept, we ran a search with *Corporate Digital Responsibility* as a single search string in the title, keywords, and abstracts (December 2024). We used inclusion criteria for language (English only) and subject categories (corporate context, economics, IT, sustainability, business ethics) and focused on peer-reviewed publications. Accordingly, we extracted 133 articles from two

databases (44 from Web of Science, and 89 from Scopus), of which 41 duplicates were dropped. We screened the remaining records and further limited the search to journal and conference articles (including editorials), dropping 16 books or book chapters and three additionally identified duplicates. After another round of close-up independent screening, we dropped 14 more articles that did not fit the research question or advance the discussion on CDR. Moreover, 10 literature reviews were excluded, and one article for which we did not have full access, even after requesting it. Finally, the remaining articles were thoroughly screened, and seven relevant articles mentioned in the references of the identified articles were further included. This resulted in a total of 55 studies (Figure 1). For an overview of all the included articles see Appendix A2 (p. 210).

Second, we searched the same databases using the search string *Corporate Social Responsibility* AND *digital**[19] (January 2025), and the same inclusion criteria described above, with a search from 2019 onwards to have a comparable timeframe with CDR articles. This initially extracted 192 articles from the two databases (92 from Web of Science, and 100 from Scopus), of which 84 duplicates were dropped, and a further six that were duplicates with the CDR literature. We excluded two review articles, two no-access articles, and 23 articles that did not fit the criterion of lying at the intersection of CSR and digitalisation. For instance, some articles focused exclusively either on digitalisation, or on CSR. In total, we retained 75 studies appertaining to CSR and digital* for our in-depth analysis (Figure 2). For an overview of all the included articles see Appendix A3 (p. 214).

[19] While we searched for *digital** in the title, keywords and abstracts, we reduced the search for *CSR* to be in the title, as we were interested in articles that clearly address digitalisation in the context of CSR.

Figure 1: PRISMA Search Protocol of CDR

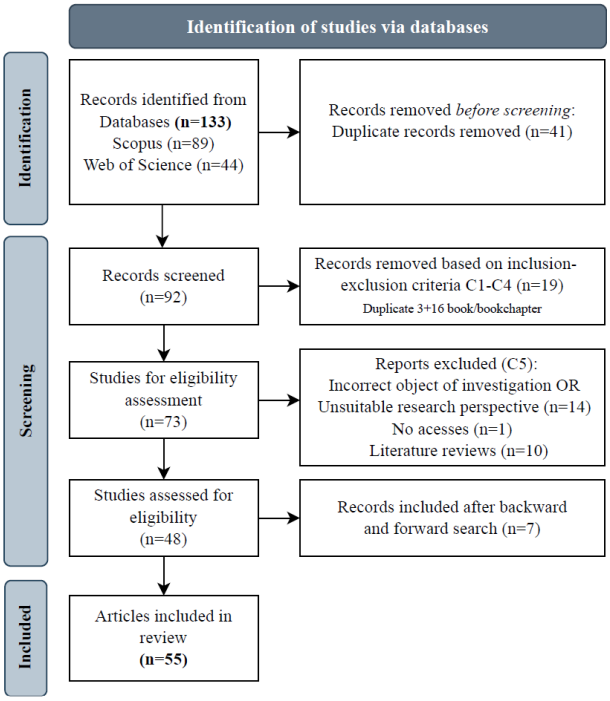
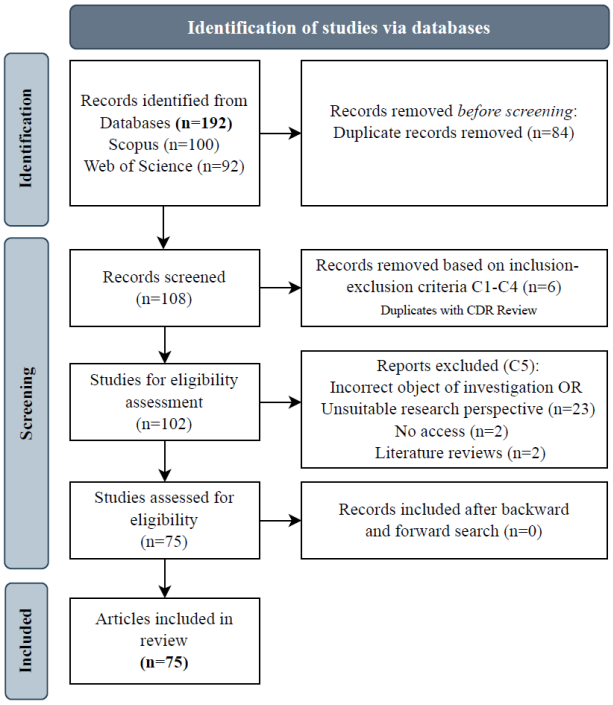


*Figure 2: PRISMA Search Protocol of CSR and Digital**



3.3 Descriptive Review

Although the research in both literature streams is flourishing, it is also widely fragmented (Figure 3 and 4). Literature about CDR has been published in 50 different outlets. The journal with the most publications is *Organizational Dynamics* (n =7). Articles on CDR follow largely theoretical and empirical qualitative methods, with an increasing tendency of quantitative and mixed methods studies in the last three years. Overall, half of the CDR articles are based on theoretical methods, and quantitative and mixed methods comprise one-quarter of all reviewed CDR articles. Moreover, the articles have a predominantly European perspective, and refer to EU regulations (e.g., EU AI Act, GDPR), first and foremost studies conducted in a German setting researchers (e.g., Merbecks, 2024; Pelters, 2021; Weber-Lewerenz, 2021).

Literature on the intersection of CSR and digital* is also fragmented across 52 different outlets, with the most hits found in the Journals *Sustainability* (n=9) and *Corporate Social Responsibility and Environmental Management* (n=5). In contrast with the theoretical focus of the CDR literature, the articles about CSR primarily (i.e., more than two-thirds) tend to apply quantitative methods. However, in contrast with the Eurocentric perspective of CDR articles, the CSR studies mainly refer to China and other

Asian countries, such as Thailand or Vietnam. Moreover, the two literature streams vary in terms of their research discipline (Table A3 and A4). CDR is more prominent in business ethics and information system research, often resulting in interdisciplinary discussions, and reflecting on aspects of organisational ethics of digital technologies (Carl & Hinz, 2024; Mueller, 2022; Stahl, 2024; Trier et al., 2023). By contrast, CSR and digitalisation are mostly located in business ethics and management

Figure 3: CDR Publications until 2024 (n=55)

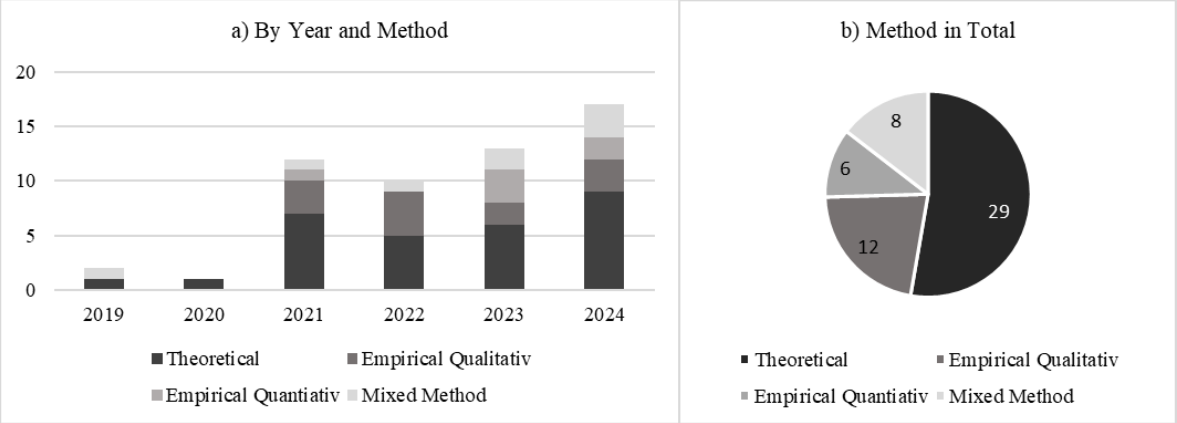
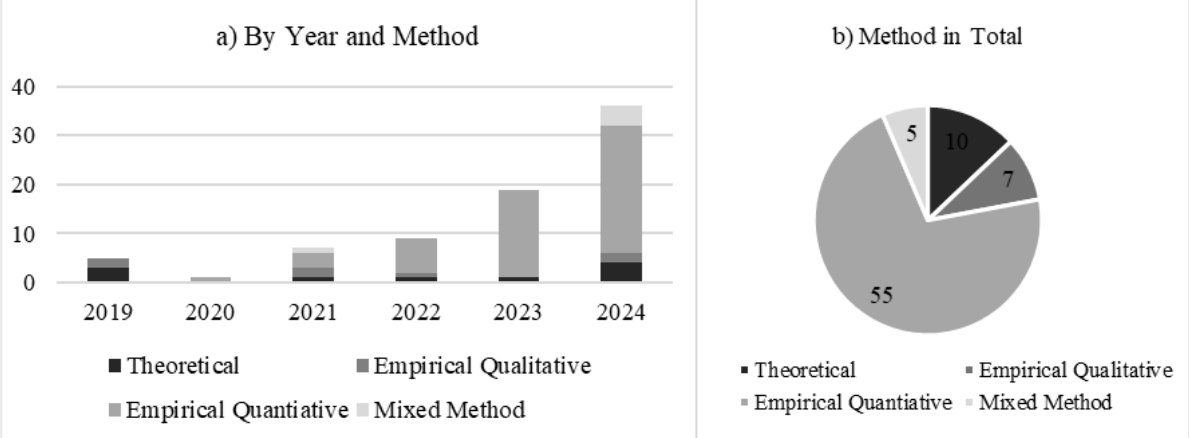
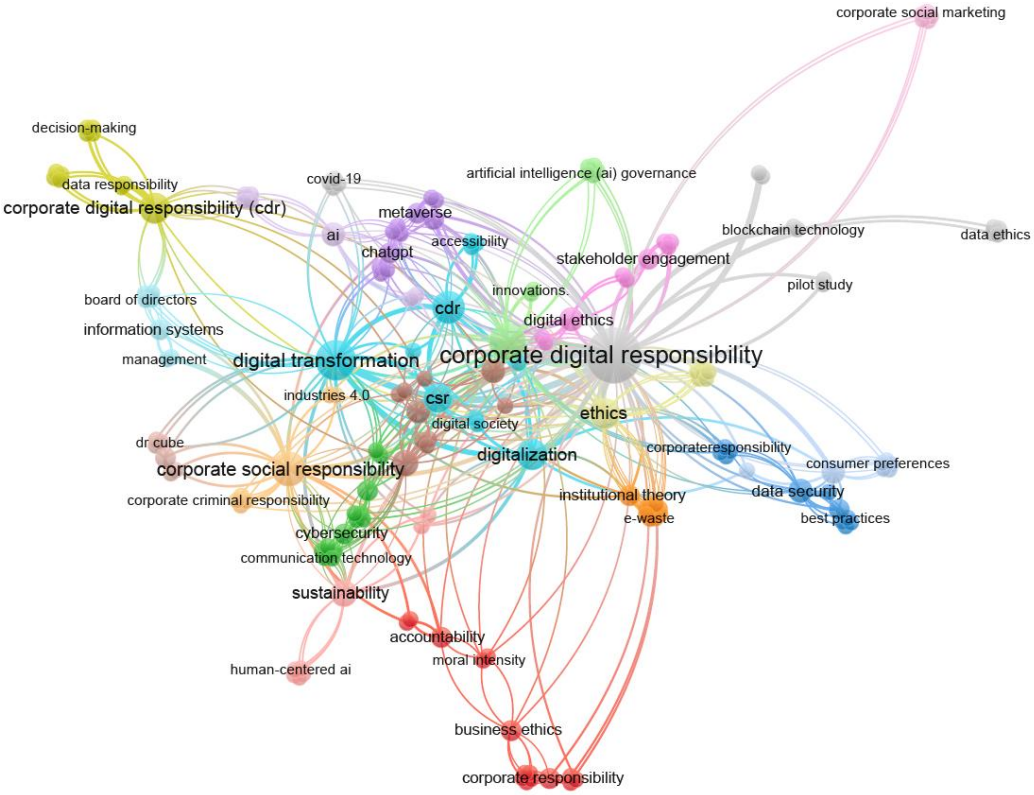


Figure 4: CSR and Digital Publications between 2019 and 2024 (n=75)*



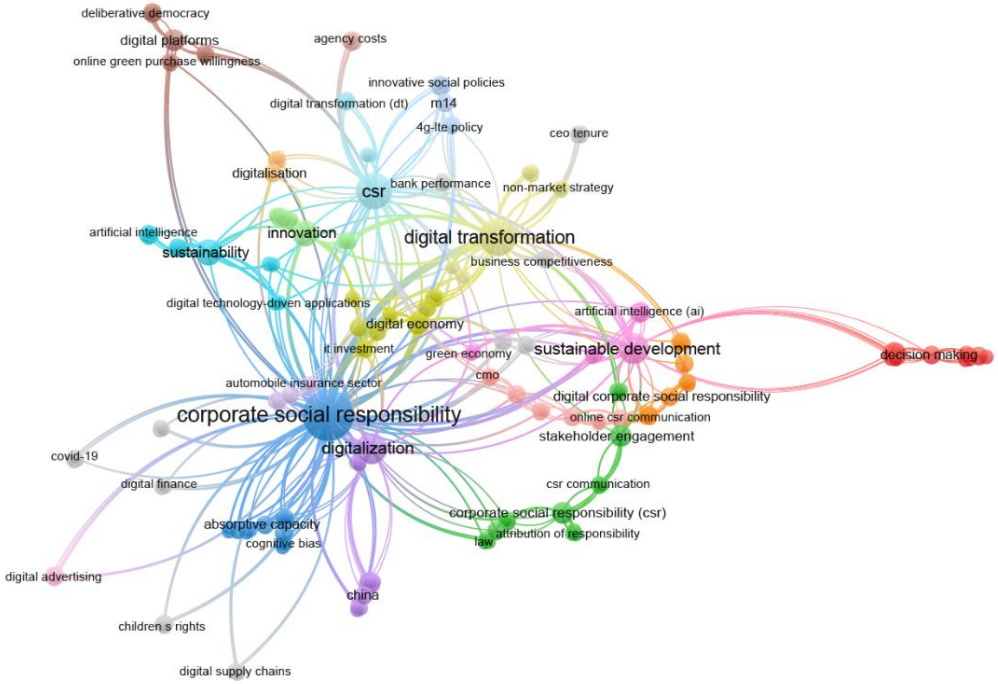
Illustrating the keywords in the two literature streams with VOSviewer makes apparent the thematical differences between them (Figure 5 und 6). The literature about CDR uses keywords like ‘CDR’, ‘CSR’, ‘digitalisation’, ‘ethics’, ‘data security’, and ‘AI governance’ more frequently, while the articles of the search string of CSR and digital* commonly uses keywords like ‘sustainable development’, ‘digital transformations’ or ‘digital platforms’ (Figure 4). The differences in their respective keywords already suggests that the literature streams of CDR and CSR tend to focus on different topics.

Figure 5: Visualisation of the CDR keyword network between 2019-24 (VOSviewer)



Note: The keywords displayed are exactly as mentioned by the articles. This means there are synonyms or abbreviations, such as CDR and Corporate Digital Responsibility.

Figure 6: Visualisation of the CSR and Digital keyword network between 2019-24 (VOSviewer)*



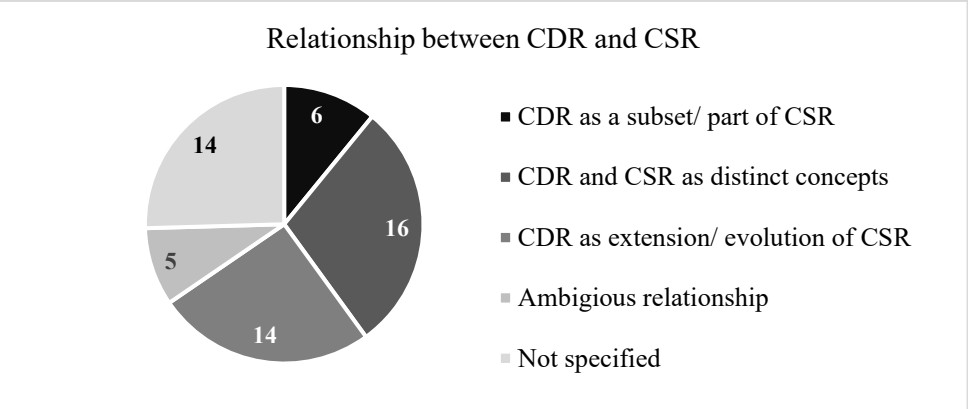
Note: The keywords displayed are exactly as mentioned by the articles. Hence, there are synonyms or abbreviations, such as CSR and Corporate Social Responsibility.

3.4 Thematic Review Results

In line with other literature reviews, we conducted a thematic coding of the identified articles concerning CDR and CSR/digital* (e.g., Schlütter et al., 2024; Thorpe et al., 2005). To understand the connection between CDR and CSR, we first identified the understanding of CDR in relation to CSR as depicted in the CDR literature. In parallel, we identified the main topics of each literature stream, with two authors coding all the articles independently and in an iterative process according to their main topics, focus, and contribution. Then, both authors jointly discussed the codes and identified the key topics of each article. This identified three overarching topics for each literature stream (see Tables in Appendix A3 and A4). Afterwards, we compared the coded topics to identify differences and commonalities.

As there are varying understandings of CDR (overview of definitions see Appendix A4, p. 219), there are also varying understandings of its relationship to CSR (Figure 7). On the one hand, scholars consider CDR to be a distinct and independent concept but acknowledge inevitable overlaps with CSR (e.g., Lobschat et al., 2021; Napoli, 2023). In another group of articles, the authors depict CDR as an extension or evolution of CSR (e.g., Clausen et al., 2023; Girrbach, 2021; Herden et al., 2021; Volchek et al., 2024). A smaller number of scholars describe CDR as a subset or cross-sectional field of CSR, hence as a part of CSR (e.g., Bernini et al., 2024; Schrödter & Weißenberger, 2024). Further, a considerable number of articles either presents the relationship of CDR and CSR as ambiguous or does not specify it at all. Mihale-Wilson et al. (2022) and Carl et al. (2023) claim that CDR and CDR conceptualisations increasingly converge, while Mueller (2022) suggests viewing CDR as a ‘transient phenomenon’, believing that the different camps will converge their viewpoints in the long term. In general, while the proposed categorisation of different understandings is not clear-cut (e.g., overlaps between CDR as an evolution of CSR or as a subset of CSR), this overview suggests that CDR has at least some distinct attributes compared to CSR.

Figure 7: Relationship between CDR and CSR presented in the literature about CDR (n=55)



Topics in the CDR Literature

We identified the following core topics in the CDR literature: (a) conceptualisations of CDR, (b) the role of Artificial Intelligence (AI) and human-centric approaches, and (c) data privacy and digital trust (Table 1).[20]

Conceptualisations of CDR

This field of research explores the definitions, principles, and practical implications of Corporate Digital Responsibility. **S**tudies on CDR often focus on foundational elements of the concept. For example, Lobschat et al. (2021) propose a framework that identifies four key stakeholder groups that corporations have to consider in their CDR decisions (i.e., institutional, governmental and legal actors; organizations; individual actors; and artificial and technological actors), and three layers of CDR culture: shared values, specific norms, and artifacts and behaviours linked to digital responsibility. This seminal work has inspired subsequent studies that explore CDR norms, principles, drivers, and outcomes (e.g., Cheng & Zhang, 2023; Vo Thai et al., 2024; Wynn & Jones, 2023). While some scholars emphasize transparency and participatory mechanisms, particularly regarding consumer privacy protection (Carl, 2021; Carl et al., 2022, 2024), others broaden the discussion to include environmental sustainability and societal well-being (Elliott & Copilah-Ali, 2024b). Digital well-being is increasingly highlighted as a dimension of organizational attractiveness (Clausen et al,. 2023), yet relatively few studies explicitly address environmental aspects, such as e-waste management, or board-

[20] These three topics are not mutually exclusive. For instance, articles discussing constitutions of CDR also partly refer to AI or to the need to gain stakeholders' digital trust. Moreover, these CDR topics also feature in CSR literature, albeit to a lesser extent.

level influences on responsible digital practices (Nagano, 2023; Napoli, 2023). Empirical evidence, such as the analysis by Merbecks (2024) of German DAX 30 companies, underscores the role of CDR as a valuable internal management tool that often features in corporate non-financial reports.

Beyond these foundational debates, some scholars advocate expanding CDR into broader theoretical perspectives, or dividing it into subcategories. The lack of a unified scientific direction remains a challenge (e.g., Volkov & Sidorenko, 2022), prompting novel concepts such as Societal CDR. Societal CDR integrates the measurement of societal impact and stakeholder theory to extend the concept beyond corporate boundaries by particularly addressing passive stakeholder groups, that are only indirectly related to the business, however, by the influenced by digitalisation (Dörr & Lautermann, 2024). Others position CDR at the intersection of corporate social responsibility (CSR) and information systems management (Volchek et al., 2024). Taxonomies have also emerged to differentiate digital responsibility at various levels, such as Trier et al.'s (2023) framework distinguishing between Personal (PDR), Corporate (CDR), and Societal (SDR) Digital Responsibility. Another perspective differentiates between corporate digitized responsibility – concerned with unbiased data collection, protection, and maintenance – and corporate digitalized responsibility, which focuses on ethical data interpretation and managing value conflicts in data-driven decision-making (Cheng & Zhang, 2023). In contrast, Stahl (2024) challenges these segmented approaches, arguing that CDR should be seen merely as a subset of broader digital ethics, and proposes a shift towards the notion of responsible digital ecosystems that transcend organizational boundaries. The field’s expanding scope is further illustrated with industry-specific applications of CDR, such as its role in blockchain-based supply chain management (Girrbach, 2021), or implementation strategies in service firms (Kunz & Wirtz, 2024).

The motivation to implement CDR is frequently linked to its potential business benefits. Maintaining a competitive advantage is a recurring justification for investing in CDR practices (e.g., Orbik & Zozuľaková, 2019; Schrödter & Weißenberger, 2024). While empirical research indicates a positive relationship between responsible digital practices and financial performance (Schrödter & Weißenberger, 2024), challenges still remain. However, challenges still remain, for instance, high development and operational costs often lead to compromises in CDR implementation (Wirtz et al., 2023). The widely cited cost-benefit framework developed by Wirtz et al. (2023) underscores the trade-offs

between opportunity costs – such as reduced personalization, less effective cross-selling, and lower supply chain optimization – and long-term benefits like trust, customer loyalty, risk mitigation, and reputational gains.

Table 1: Identified Topics in CDR and Examples of Research Objectives

Conceptualisations of CDR Topics	
Conceptualisation	• What is the state of research on CDR? (Paluch, 2024) • How can Societal CDR give CDR a broader societal perspective? (Dörr & Lautermann, 2024) • The lack of a unified scientific direction for corporate responsibility on digital platforms (Volkov & Sidorenko, 2022) • Distinguishing *digitized* responsibility from *digitalized* responsibility (Cheng & Zhang, 2023)
Norms, principles, drivers, outcomes	• How organizations make responsible strategic decision (…) under the framework of CDR. (Rugeviciute, 2024) • To develop recommendations for a corporate digital responsibility (CDR) strategy (Volchek, 2024) • CDR can be interpreted as one facet of digital ethics that focuses on how topics and questions from digital ethics are perceived and dealt with by organisations (Stahl, 2024) • CDR codifies TRUST and illustrates how AI governance and expectations are met, building on lessons learned from CSR (Elliott et al., 2021)
Debate about affiliations with CSR	• CDR and CSR represent complementary but also sometimes overlapping concepts of business ethics (Carl, 2021) • While CSR aims to minimize the negative impacts, and maximize the positive outcomes of corporate practices on socially and environmentally relevant issues, CDR intends to minimize the adverse effects of digitization while maximizing the positive impacts of corporate digital activities (Carl, 2021) • Why do we need CDR in addition to CSR? Does the established CSR concept not cover CDR too? (Mihale-Wilson et al., 2022)
Modelling and measuring CDR	• How can CDR and its parameters be modelled more simply? (Wynn & Jones, 2023) • How can CDR be conceptualized, measured, and implemented? (Cheng & Zhang, 2023) • How do institutional entrepreneurs understand/use CDR for responsible digital innovation? (Trittin-Ulbrich & Böckel, 2022)
Role of Artificial Intelligence and a Human-centric Approach	
Potential adverse effects of AI	• Measuring machinewashing as deceptive communication in the context of AI (Bernini et al., 2024) • How can digital societal harms be avoided in AI systems using CDR? If we permit AI to make life-changing decisions, what are the opportunity costs, data trade-offs, and implications for social, economic, technical, legal, and environmental systems? (Elliott et al., 2021) • Threats posed by the development of new technologies, artificial intelligence, automation, and digitalisation of the social environment on a large scale (Suchacka, 2019) • The increasing use of AI and digital technologies has led organizations to face complex ethical, fairness, and privacy challenges. In addressing these concerns, the concept of CDR has been introduced (Hartley et al., 2024)
Specific AI applications	• How can CDR be applied in AI retail service? (Scarpi & Pantano, 2024) • How can the AI accountability framework and CDR be implemented in financial services? (Tóth & Blut, 2024)

	• How will intelligent automation, service robots, and AI reshape service products and their delivery, particularly focusing on the implications for service firms and their marketing strategies? (Wirtz & Pitardi, 2023) • Various cases of AI technology applications underscore the urgency of CDR (Paluch et al., 2024)
Navigating AI & human-centric approaches	• How are organizational sensemaking processes of creation, interpretation, and enactment triggered by conversational AI issues and events? (Sidaoui et al., 2024) • How should an adequate ethical framework be designed to support digital innovations in order to make full use of the potentials of digitization and AI? (Weber-Lewerenz, 2021) • What is CDR and how do advances in AI affect it? (Kunz & Wirtz, 2024) • CDR is a potential framework to assist navigating AI governance complexity and to devise an informed strategy (Elliott et al., 2021) • The dynamic development of AI has also accentuated the pressing need for corporate CDR, with special regards to identifying accountability and human agency (Tóth & Blut, 2024) • Managers should adopt frameworks that help in navigating the ethical challenges associated with AI, such as managing customer vulnerabilities and data usage (Paluch et al., 2024)
Data Privacy and Digital Trust	
Governance & data privacy	• How do companies incorporate digital compliance as part of CDR? (Schrödter & Weißenberger, 2024) • How can CDR activities be evaluated at the company level, particularly focusing on privacy and data security? (Carl et al., 2022) • How can companies ethically communicate their data privacy and security practices in the context of evolving consumer expectations and responsibilities? (Carl, 2022)
Specific industry	• How can consumer data vulnerability in online banking be minimized by market-oriented CDR? (Liyanaarachchi et al., 2021) • Investigating the accountability gap in business-to-government data sharing, and how CDR can fill this gap (Schneider, 2022) • Access to digital data in the recruitment process (Peshkova, 2022) • How can managers, especially in the retail and advertising sectors, ensure that AI implementations respect data privacy and promote ethical standards? (Paluch et al., 2024)
Digital trust	• How can CSR and CDR mechanisms be used for implementing responsible data use? (van der Merwe, 2022) • How does CDR influence digital trust? (Jelovac et al., 2022) • How can firms measure Digital Trust? (Kluiters et al., 2023)

Role of Artificial Intelligence and a Human-centric Approach

Second, CDR consolidates discussions on the implications and challenges of AI and the need to refocus on a human-centric approach. Surprisingly, although AI was mentioned numerous times, it was seldom defined. Admittedly, defining AI is a thankless task, as there is no common ground among users (Stahl, 2024)[21], beyond its general

[21] Stahl (2024) further states that 'One of the reasons for the concerns is that many instantiations of current machine learning technologies are difficult or impossible to fully understand, even by the experts who build them'. Pappas et al. (2023) point to numerous discussions in the IS literature, focusing on AI, Responsible AI, Explainable AI, Human-centric AI, or inclusive AI.

description as scaled computable machine learning algorithms (Elliott et al., 2021). Weber-Lewerenz (2021) stated that the term was first coined by John McCarthy in 1955, and its ethical challenges are nowadays reflected in the interdisciplinary debate 'Ethics in AI'. Some CDR scholars refer to the adverse effects of AI and particularly generative AI (GenAI)[22], citing threats to the labour market and to human-machine relationships (e.g., Suchacka, 2019, 2020), critical changes in the workforce (Kunz & Wirtz, 2024), its black box character (Elliott et al., 2021), complex ethical, fairness, and privacy challenges (Hartley et al., 2024), or the need to mitigate biases (Wynn & Jones, 2023). CDR is introduced as a remedy to address and govern these challenges and to use AI to its full potential, such as for improved problem-solving, increased productivity, and efficiency (e.g., Elliott et al., 2021; Hartley et al., 2024). Robot- and AI-powered services are expected to improve customer experiences if responsibly handled, where CDR may enable the identification of accountability (Tóth & Blut, 2024; Wirtz & Pitardi, 2023).

CDR is a potential framework to navigate the complexities of AI governance, refocus on a human-centric view, and allow developers and engineers to reflect on the ethics by design (Elliott et al., 2021; Elliott & Copilah-Ali, 2024b; Weber-Lewerenz, 2021). Elliott et al. (2021) have found more than 160 AI principles advocating corporations to act responsibly. They further highlight that '[t]his maelstrom of guidance, none of which is compulsory, serves to confuse, as opposed to guide' and state that 'CDR [is] a potential collaborative mechanism to demystify governance complexity and to establish an equitable digital society ' (Elliott et al., 2021: 179). The authors point to the Draft EU AI Regulation (DEAR) of 2021, but the latest developments include the EU 'AI Act'[23] from 2024, a legal framework to guide AI developers, deployers, and users providing classifications of risk levels.

Several studies reflect on CDR in specific AI applications. For instance, Scarpi and Pantano (2024) discuss the role of CDR in AI for intelligent retail service by developing the new concept of Artificial Intelligence Responsibility in Retail Service Automation (AIRRSA). Pappas et al. (2023) provide an overview of human-AI partnerships and responsible AI of industry 5.0 for a sustainable society, whereas Paluch

[22] GenAI is described as "a subset of AI that involves the use of algorithms to generate human-like outputs, whether it is text, images, voice, or other forms of content" (Sidaoui et al., 2024).
[23] For more information see DEAR: https://eur-lex.europa.eu/procedure/FI/2021_106; the EU approach https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence; "AI Act": https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng?utm_source=chatgpt.com.

et al. (2024) reflect on CDR studies addressing AI and digitalisation. Sidaoui et al. (2024) investigate the responsible integration of conversational agents (CAs) like chatbots. The authors map CDR factors such as CDR culture, CDR management structure, and digital governance, and consider the role of generative AI (GenAI) for service firms, software providers, and customers/society. Moreover, artificial actors are regarded as a specific type of indirect stakeholder that increasingly interacts with us while reducing human interactions with consumers (Kunz & Wirtz, 2024).

While many articles emphasise the need to reflect on the consequences of AI, concrete measures that companies can take, and how responsibility can be exercised, are still emergent topics in CDR literature.

Data Privacy and Digital Trust

Third, many CDR studies discuss data privacy and the need to gain digital trust. Carl (2021) describes consumer privacy and data security as significant concerns linked to information technologies and systems. Even though regulations for data privacy protection have increased (Hartley et al., 2024), the majority of legal frameworks (e.g., GDPR[24]) impose only a minimal level of requirements on companies. Applying CDR norms as a voluntary obligation could imply going beyond compliance and providing strategic initiatives to ensure greater privacy and strengthen citizen and consumer rights in making deliberate decisions upon their data (e.g., Mogaji et al., 2023). Thereby, CDR becomes a gap-filler between public law-based and data protection-law based accountability models – especially as there is 'no "invisible hand" assuring the lawful and responsible nature of occurred data processing patterns' (Schneider, 2022). Some scholars even propose CDR-related privacy and data security sub-dimensions and measurement tools (e.g., Carl, 2021; Carl et al., 2022, 2024). After all, any actions that are only symbolic but not substantial could be harmful if uncovered as greenwashing or machinewashing[25] (Bernini et al., 2024; Famularo, 2023). These issues are particularly apparent in the banking and finance sector (Liyanaarachchi et al., 2021). Schrödter and Weißenberger (2024) apply institutional theory and empirically examine how companies incorporate aspects of digital compliance to gain competitiveness and legitimacy. They

[24] General Data Protection Regulation, available at: https://t1p.de/c3on2

[25] Machinewashing is defined as a business strategy for the ethical use of AI and algorithm-based systems, based on misleading behaviour affecting reporting (omitted or misleading information provided by words and images) and/or action (the underlying algorithm of AI) directed at various critical stakeholders to gain their acceptance" (Bernini et al., 2024).

illustrate that data protection is mostly institutionalised. Kunz and Wirtz (2024) summarise the potential trade-off between organisational goals and CDR practices, as the financial benefits of data collection and customisation might prevent good CDR practices. In a similar vein, Van Der Merwe and Al Achkar (2022) state the need for data responsibility to be placed at the core of the business. This is also reflected in Peshkova (2022), who describes CDR in terms of access to the digital data of employees and even job candidates.

Data privacy and security endeavours are closely linked to corporate communication and the aim of gaining consumer trust (Carl, 2022). Trust, which is generally considered pivotal for companies' success in the digital realm (Elliott et al., 2021; Hartley et al., 2024; Jelovac et al., 2022; P. Jones & Comfort, 2021), is an often-debated outcome of CDR, and the redefinition of trust in the digital sphere seems to be particularly important (e.g., Elliott et al., 2021; Hartley et al., 2024; Kluiters et al., 2023). Jones and Comfort (2021) outline trust in maintaining customer confidence about digital information privacy in sports betting. Digital trust differs from personal peer-to-peer trust and arises from challenges, including data privacy and protection (e.g., Stahl, 2024). Kluiters et al. (2023) describes digital trust as 'represent[ing] stakeholders' confidence in the competence of actors, technologies, platforms, and processes of establishing a reliable network', thereby complementing and partly even replacing personal peer-to-peer trust Kluiters et al. (2023:74). They further stress that customers have higher demands for platforms than for in-person interactions. Therefore, findings of CSR and trust cannot easily be transferred to the digital sphere. It is argued that, as CDR positively affects digital trust and customer loyalty, it can attract ethically conscious consumers (e.g., Dörr & Lautermann, 2024; Hartley et al., 2024; Kärpänen, 2022).

Topics in the Literature intersecting CSR and Digitalisation

For the Corporate Social Responsibility and digital* literature we followed the same process as we did for the CDR literature. After searching for keywords, and both authors independent coding the articles, we agreed on the following three overarching topics (Table 2): (1) how digitalisation and CSR interact, (2) how CSR is connected to performance in the digital context, and (3) how CSR is influencing/influenced by the digital communication context.

The Interaction between Digitalisation and CSR

The scholarly discussion about this interaction is articulated around three main lines of thinking. In the first, scholars explore how digital technologies, such as AI or the Internet of Things, have been transforming CSR (Govindan, 2024; Shestakova, 2024; Shkalenko & Nazarenko, 2024). Some studies investigate the integration of CSR in specific digital contexts (Etter et al., 2019; Gilbert et al., 2024; Guo et al., 2024; Lindman et al., 2023; H. Wang et al., 2024). Other studies discuss the role of big digital platforms (e.g., Google, Apple, Meta) as gatekeepers and in the implementation of CSR (Gilbert et al., 2024; Lindman et al., 2023; Y. Ma et al., 2024). Lindman et al. (2023), for instance, argues that the political power of digital platforms requires a different understanding of CSR, while Gilbert et al. (2024) consider how digital platforms can exhibit ethical responsibility towards stakeholders.

Second, while the previous discussion considers how CSR itself is affected by the digital context, another debate centres around how digitalisation is a driver for CSR. This literature mostly assumes that digitalisation positively influences or even enables CSR. Researchers analyse how different forms of digitalisation or digital technologies—including digital transformation, AI, and blockchain—enhance CSR. (W. Chang et al., 2023; L. Chen & Chen, 2023; González-Ramírez et al., 2024; Kong & Liu, 2023; Xin et al., 2022; Zhou et al., 2024). Industry-specific studies explore how digitalisation impacts CSR in fintech, mining, insurance, and agriculture (Abad-Segura et al., 2024; Abdallah-Ou-Moussa et al., 2024; B. Li et al., 2024; Y. Xu et al., 2023). Additionally, one study examines the effects of digitalisation – techno-invasion[26], especially – on stakeholders' well-being (Aleksić et al., 2024). Some scholars investigate the mechanisms linking digitalisation with CSR, such as the mediating role of innovation (Jiang et al., 2023). Finally, some studies explore other aspects of the relationship between digitalisation and CSR, such as the influence of CSR on digitalisation (Djakman & Siregar, 2024; Stock et al., 2022; Zhu et al., 2024), the bidirectional relationship between digitalisation and CSR (Jiang et al., 2023), and digitalisation as a moderator, which amplifies the impact of CSR (R. Huang & Wei, 2023).

[26] The authors refer to techno-invasion as a "dimension of technostress referring to being constantly connected and thereby invading the employee's personal life " (Aleksić et al., 2024:430).

Table 2: Identified Topics in CSR and Digitalisation and Examples of Research Objectives

The Interaction of Digitalisation and CSR	
How digitalisation changes CSR	• How digitalisation is changing CSR (Govindan, 2024; Shestakova, 2024; Shkalenko & Nazarenko, 2024) • How digitalisation and CSR are intertwined in specific contexts such as platforms or the sharing economy (e.g., How can digital platforms exhibit CSR?) (Etter et al., 2019; Gilbert et al., 2024; Guo et al., 2024; Lindman et al., 2023; M. Wang et al., 2024)
The influence of digitalisation on CSR	• Effects of different forms of digitalisation on CSR, generally (e.g., digital technologies, digital transformation) (W. Chang et al., 2023; L. Chen & Chen, 2023; González-Ramírez et al., 2024; Kong & Liu, 2023; Xin et al., 2022; Zhou et al., 2024) • Effects of digitalisation on CSR in specific industries such as fintech, mining, insurance, agriculture (Abad-Segura et al., 2024; Abdallah-Ou-Moussa et al., 2024; B. Li et al., 2024; Y. Xu et al., 2023) • Effects concerning specific stakeholders (e.g., employees) (Aleksić et al., 2024) • The mechanisms that connect digitalisation with CSR (Jiang et al., 2023)
Further relationships	• The influence of CSR on digitalisation (Djakman & Siregar, 2024; Stock et al., 2022; Zhu et al., 2024) • The bidirectional relationship between digitalisation and CSR (G. Huang & Shen, 2024) • Digitalisation as an enhancing moderator effect (R. Huang & Wei, 2023)
Performance Perspective	
The interactive effect of digitalisation and CSR on organisational performance	• How integrating sustainable technological innovation with CSR can improve a company's competitiveness (Abad-Segura et al., 2024) • The relationship between digitalization investments, CSR, and bank performance (Thuong, 2024) • How digital transformation positively enhances the relationship between CSR and organizational resilience (H. Wang et al., 2024) • The effect of digital transformation on innovation performance with CSR acting as a moderating factor (L. Wang & Yan, 2023) • How internal CSR reduces the negative impact of digitalization on firms' innovation efficiency and how internal corporate social irresponsibility (CSiR) strengthens this negative effect (Zhong & Ren, 2024) • The impact of digital economy development and CSR on promoting low-carbon innovation (W. Chen, 2023) • How digital platform capability enhances the positive association between CSR fulfilment and business model innovation (Liao et al., 2023)
The effect of different forms of digitalisation on CSR/ sustainability performance	• How digital CSR and digital social responsibility influence sustainability (K. S. Al-Omoush, 2024; Khattak & Yousaf, 2022) • How digital technologies improve sustainable firm performance (S. A. R. Khan et al., 2024) • How CSR enhances the effect of intelligent automation on sustainability performance(Ghobakhloo, Asadi, et al., 2023b) • How digital economy, organisational digitalisation, and digital innovation affect CSR performance (Hu & Liu, 2023; Jiang et al., 2023; H. Li et al., 2024; H. Liu et al., 2024; Srivetbodee & Igel, 2021; Sun et al., 2024; Q. Yang & Jin, 2024; L. J. Zheng et al., 2023b)
The effect of digital CSR on organisational performance	• No effect of digital CSR on ROE, ROA, but on turnover (Almeida et al., 2022) • Positive effect of digital CSR on competitive intelligence, organizational resilience, social entrepreneurship (K. Al-Omoush et al., 2024) • Digital supply chain announcements disclosing CSR information generate positive market reactions and positive impact on stock market (W. Liu et al., 2024)

The mediating role of CSR	• How CSR mediates the relationship between information technology and company's overall financial performance (Alfalah et al., 2022) • How CSR mediates the relationship between digital transformation and green technology innovation (Y. Zheng & Zhang, 2023) • How CSR mediates the relationship between information technology investment and firm performance (Alfalah et al., 2022)
The mediating role of digitalisation	• How digital transformation mediates the relationship between CSR and firm innovation (Tuyen et al., 2023) • How digital transformation mediates the relationship between CSR in technological innovation and sustainable competitive performance (Wu et al., 2024)
Effects of and on CSR in a Digital Communication Context	
Effects on perceptions and behaviours of stakeholders	• How CSR affects consumers in specific digital contexts such as e-commerce, food delivery, or museums (Fu et al., 2023; Prisco et al., 2024; Q. Shen et al., 2024) • How digitalisation influences stakeholder perceptions of CSR (Esposito & Ricci, 2021) • How CSR and corporate social irresponsibility influence specific consumer reactions such as electronic word-of-mouth or corporate reputation (Jung et al., 2022; R. Ma et al., 2021; Peng et al., 2024)
The role of stakeholder engagement	• How digital CSR communication does not necessarily lead to customer engagement (Okazaki et al., 2020) • How customer engagement mediates the relationship between digital social responsibility and CSR (Khattak & Yousaf, 2022) • How organisation engage stakeholders in their CSR programs to influence attitudes and perceptions of stakeholders (H. Park et al., 2021) • How information systems influence user engagement enabled by CSR (Zhou et al., 2024)
Digital corporate communication and CSR	• How stakeholder salience changes in view of online CSR reviews (Bhattacharyya, 2023) • How big platforms use CSR to legitimize their market expansion (Campoamor, 2019) • How CSR is portrayed in the generalist digital press (Arnal-Pastor & Berné-Martínez, 2024) • The role of codes of conduct in digital communication (López Jiménez et al., 2021) • How chief marketing officers communicate CSR online (Özturan & Grinstein, 2022)
(Ir)responsible online communication	• How unfair commercial practices are related to CSR (Vítová, 2022) • How digital image manipulation and CSR are linked to stakeholder health (McBride et al., 2019)

The Performance Perspective

Second, we found studies that explore the role of digitalisation and CSR in relation to organisational performance. One area of research focuses on the interactive effect of digitalisation and CSR on organisational performance, considering organisational outcomes such as financial performance, effectiveness, and resilience (Abad-Segura et al., 2024; Thuong, 2024; H. Wang et al., 2024). Some more practically oriented studies focus on the interaction between digitalisation and CSR with innovation, its performance

and efficiency (W. Chen, 2023; Liao et al., 2023; L. Wang & Yan, 2023; Zhong & Ren, 2024). A related area examines the effect of different forms of digitalisation on CSR and sustainability performance. Adopting an overall positive standpoint, scholars investigate how CSR enhances the effect of digitalisation on CSR performance (Hu & Liu, 2023; Jiang et al., 2023; H. Li et al., 2024; W. Liu et al., 2024; Srivetbodee & Igel, 2021; Sun et al., 2024; Q. Yang & Jin, 2024; L. J. Zheng et al., 2023). Other studies adopt the same perspective to analyse the effects on sustainability performance (K. S. Al-Omoush, 2024; Ghobakhloo, Asadi, et al., 2023b; S. A. R. Khan et al., 2024; Khattak & Yousaf, 2022).

Another research stream considers the direct effects of digital CSR on organisational performance, with some studies suggesting that, while digital CSR has little impact on financial performance, like return on equity (ROE) and return on assets (ROA), it has a positive influence on turnover (Almeida et al., 2022). Market-focused studies find that CSR-related digital supply chain announcements generate positive investor reactions and stock market gains (W. Liu et al., 2024). Others argue that digital CSR contributes to competitive intelligence, organisational resilience, and social entrepreneurship (K. Al-Omoush et al., 2024). Another stream of research focuses on CSR as a mediator between digitalisation and performance. Scholars explore how CSR mediates the relationship between IT investments and financial performance (Alfalah et al., 2022), and how CSR links digital transformation to green technology innovation (Y. Zheng & Zhang, 2023). Conversely, some studies investigate the mediating role of digitalisation, showing that digital transformation can mediate the relationship between CSR and firm innovation (Tuyen et al., 2023), and between CSR-driven technological innovation and sustainable competitive performance (Wu et al., 2024).

Effects of and on CSR in a Digital Communication Context

Third, another dominant research topic examines how CSR is influencing/influenced by a digital communication context, focusing mainly on consumer-related contexts. This topic area includes studies analysing how digitalisation affects stakeholder perceptions and behaviours toward CSR. Studies investigate CSR's impact on consumer behaviour in digital contexts such as e-commerce, food delivery, and museums (Fu et al., 2023; Prisco et al., 2024; Q. Shen et al., 2024). Other research explores how digitalisation shapes stakeholder perceptions of CSR, emphasising the role of digital transparency in shaping public trust (Esposito & Ricci, 2021). Additionally,

scholars assess how corporate social irresponsibility influences consumer reactions, including electronic word-of-mouth and corporate reputation (Jung et al., 2022; M. Li, 2021; R. Ma et al., 2021; Peng et al., 2024). Another research focus is the role of stakeholder engagement in digital CSR. Studies indicate that, while digital CSR communication is widespread, it does not always translate into customer engagement (Okazaki et al., 2020). Other scholars examine how customer engagement mediates the relationship between digital social responsibility and the effectiveness of CSR (Khattak & Yousaf, 2022), and how firms use CSR initiatives to shape stakeholder attitudes (H. Park et al., 2021). Moreover, studies explore how information systems enable user engagement in CSR initiatives (Zhou et al., 2024).

Finally, researchers investigate the effects of digital corporate communication and its relationship to CSR. Studies examine how online CSR reviews impact stakeholder salience (Bhattacharyya, 2023) and how large digital platforms use the CSR discourse to justify market expansion (Campoamor, 2019). Other research assesses how CSR is portrayed in digital media (Arnal-Pastor & Berné-Martínez, 2024), the role of corporate codes of conduct in digital communication (López Jiménez et al., 2021), and how chief marketing officers use digital platforms to communicate CSR initiatives (Özturan & Grinstein, 2022). Additionally, scholars analyse (ir)responsible digital CSR communication – e.g., how digital image manipulation in CSR messaging affects stakeholder health (McBride et al., 2019; Vítová, 2022).

Main Areas of Common Ground between the CDR and the CSR Literatures

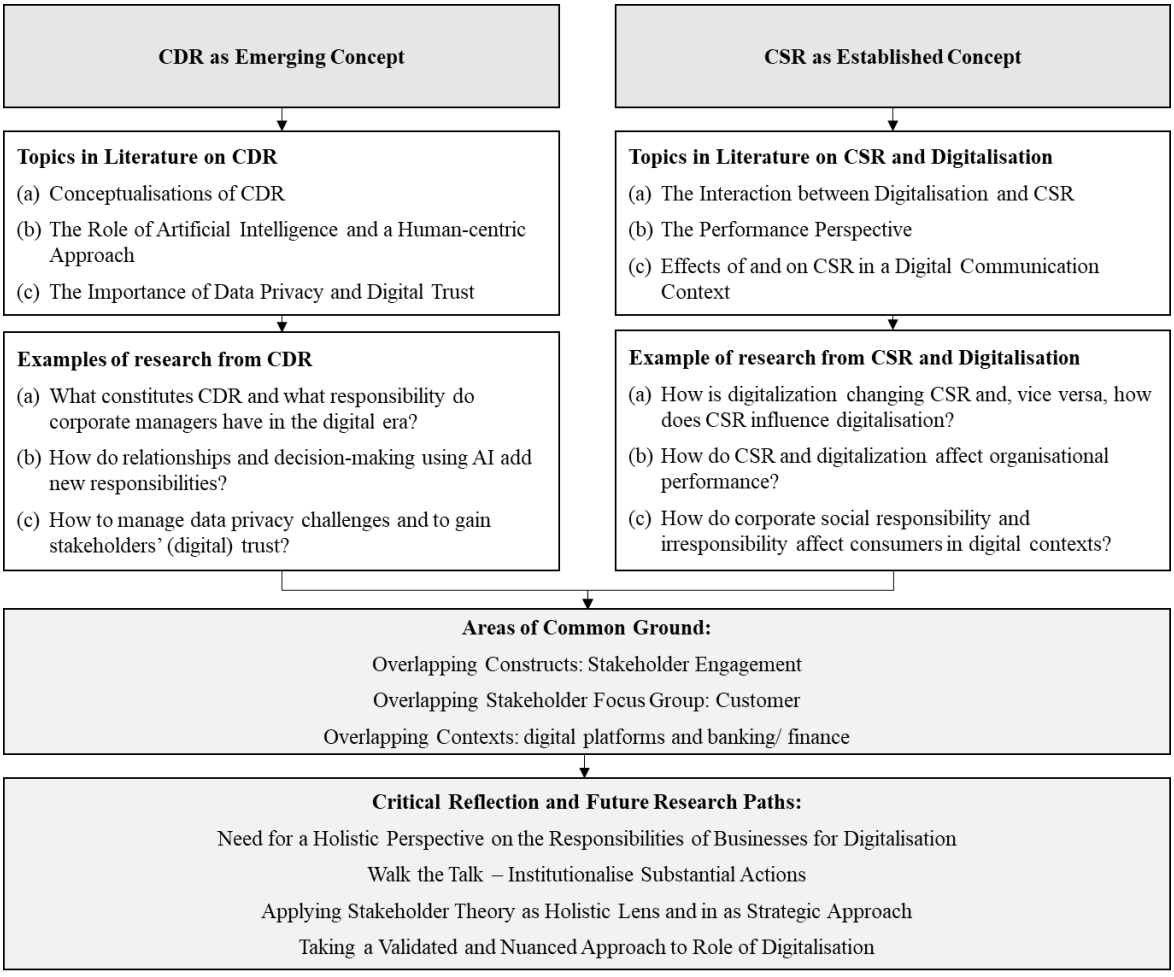
Beyond the differences in topics outlined above, both literature strands also overlap in at least three ways: they apply the stakeholder approach, they consider primarily customers, and they conduct their studies in similar contexts, like digital platforms and banking/finance (Figure 8).

Overlapping constructs

First, we noticed the frequent mention of stakeholders, and constructs like stakeholder engagement or involvement. For instance, Vo Thai et al. (2024) investigate stakeholder engagement in CDR strategy formulations, and in execution and firm performance, while Bernini et al. (2024) point to the need to empower engagement processes in the stakeholder network. Elliott and Copilah-Ali (2024) stress that wicked problems, like the

challenges induced by digitalisation and sustainability, can only be approached by 'real stakeholder involvement'. Meanwhile, Trittin-Ulbrich and Böckel (2022) refer to multistakeholder efforts that shape a firm's CDR, which then enhances institutional entrepreneurship towards responsible digital innovation. Similarly in CSR, L. J. Zheng et al. (2023) refer to stakeholder engagement as the involvement of multiple stakeholders to achieve sustainable development goals. Park et al. (2021) investigate stakeholder engagement in CSR communications, and Esposito and Ricci (2021) in a case study on virtual museums. While constructs of the stakeholder approach appear to be similarly referred to, many articles only weakly apply these constructs explicitly. This provides a rational for converging both concepts based on stakeholder approaches as will be outlined in future research.

Figure 8: Overview of Findings in CDR and CSR intersecting Digitalisation



Overlapping Dominant Stakeholder Group

Second, many studies in both literature strands frequently mention one stakeholder group particularly, namely consumers/ customers. The CDR literature addresses customers' concerns over data privacy, or digital trust (Carl, 2021; Carl et al., 2022, 2024; Mihale-Wilson et al., 2022; Paul et al., 2024). As such, CDR scholars emphasise the idea of positioning 'vulnerable customers as a critical stakeholder' (Liyanaarachchi et al., 2021: 571). In CSR, scholars point to consumer purchase behaviours (e.g., Fu et al., 2023; Shen et al., 2024), consumer attitudes or perceptions (e.g., Khattak & Yousaf, 2022; Prisco et al., 2024; S. Yang et al., 2023), and consumer engagement or dialogue (e.g., Okazaki et al., 2020; Pan et al., 2021). Apart from customers, employees are also mentioned in CDR, albeit less often. For instance, Clausen et al. (2023) refer to the implications of knowledge work, Dörr and Lautermann (2024) to unfair working conditions of gig and click workers, and Elliott and Copilah-Ali (2024) stress the challenges for leaders, managers and employees to navigate wicked problems in digitalisation and sustainability. By contrast, employees are far less often emphasised in the CSR and digitalisation literature – for instance, when discussing the implications of digitalisation on job losses (Shestakova, 2024), or its effect on employee stress (Aleksić et al., 2024).

Overlapping Contexts

Third, both literatures overlap in at least two contexts: platform/ social media and banking/ finance, with CDR scholars discussing the advantages and ethical challenges of digital platforms. Digital platforms offer companies transparency over customer and market behaviour and enable servitization, sending signals or excluding bad actors (Kluiters et al., 2023b; Wirtz et al., 2023). On the other hand, managing the flood of data and ensuring data security can be challenging, and can create adverse consequences for social media on user well-being (Kluiters et al., 2023b; Wirtz et al., 2023). Similarly, CSR scholars explore, for instance, the implications of the sharing economy (Etter et al., 2019), digital platforms as gatekeepers and their impact on advancing moral legitimacy (Gilbert et al., 2024), consumption behaviour on platforms and the impact of CSR (Fu et al., 2023), platform CSR governance (Guo et al., 2024) and price discrimination in e-commerce platform trading (Ma et al., 2024). The context of banking is also addressed in both literatures. For instance, CDR discusses the data security of online banking

(Liyanaarachchi et al., 2021) and the ethical challenges of automated financial advice via algorithmic trading (e.g., Elliott et al., 2021; Tóth & Blut, 2024). CSR scholars explore the impact of CSR on banking, for instance, in a case study of the Taiwan Bank industry (Yang et al., 2023), or in Vietnam's Banking sector (Thuong, 2024), or more generally on the world's most digitised banks (Rangel-Pérez et al., 2023).

3.5 Critical Reflections and Future Research Paths

Synthesising and comparing the literature on CDR and CSR intersecting digitalisation has revealed some key distinctions but also some crucial areas of common ground. In the following we will outline some vital research paths, additionally illustrated in Figure 9.

Need for a Holistic Perspective on the Responsibilities of Businesses for Digitalisation

Scholars exploring how corporations can undertake digitalisation responsibly often face a challenge in deciding whether to adopt Corporate Digital Responsibility (CDR), Corporate Social Responsibility (CSR), or both. This study aims to clarify the boundaries of each concept and explore where they intersect. While both CDR and CSR are valuable in their own right, there is a need for a discussion forum to prevent confusion and ensure that research is aligned and clear. By reconciling these concepts, scholars can learn from each other, and future research can focus on developing shared, agreed-upon frameworks. However, challenges arise from having both concepts in play. Since digitalisation affects all areas of business, it becomes difficult to distinguish what aspects are purely CSR and what is influenced by digitalisation. For example, both CSR and CDR can address environmental or social harms caused by digital technologies, but which perspective should take precedence? Furthermore, digital technologies can play a critical role in accelerating CSR goals—how can these technologies be designed to follow CSR principles while minimizing unintended consequences like rebound effects? A holistic approach, considering social, economic, and environmental impacts, can help align CSR and CDR efforts in the digital transformation, promoting long-term sustainability and guiding the responsible application of digitalisation.

Walk the Talk – Institutionalise Substantial Actions

Compared to the CSR in digitalisation literature, the CDR literature tends to be more theoretically based and often includes essential ethical debates, appertaining to AI or data privacy challenges, for example. Hence, CDR discussions about ethical challenges of digitalisation seem to focus more strongly on the responsibility aspect of these concepts, that is, for whom, what, and why corporations should take responsibility in the digital age. However, it becomes important to walk the talk in the sense that ethical statements need to be backed up with substantial, measurable, and tangible practices to avoid the risk of being perceived as machinewashing. As the literature on CSR and digitalisation tends to focus on empirical investigations, both literatures can learn from each other and offer a more holistic perspective on normative discussions and measurable practices.

After all, the increase in regulations such as the EU AI Act and GPDR raises the question of what is left for corporations to take on as voluntary commitments. How much regulation is needed, and can society rely on corporate self-commitment? How do corporations adapt local responsibility norms, and what does responsibility entail instead of legal obligations? Theoretical underpinnings such as social contract theory, stakeholder theory, or signalling theory may outline the advantages of voluntary self-commitments, and answer the question of how to hold companies accountable. These discussions could expand the discussion of the CDR cost-benefit calculus. After all, addressing and balancing short- and long-term costs and benefits would require anchoring CDR and CSR in corporate strategy.

Applying Stakeholder Theory as a Holistic Lens and with a Strategic Approach

Following a strategic perspective, a promising approach is stakeholder theory. After all, digitalisation involves constantly changing stakeholder relationships and expectations. For instance, if artificial actors are considered a stakeholder group, this would question the typical peer-to-peer relationships in stakeholder theory, where stakeholders have names and faces (McVea & Freeman, 2005). Currently, the literature on CDR and CSR in digitalisation primarily focuses on customers. While several studies frequently mention stakeholders, we want to highlight that the term ‘stakeholder’ itself is generally a contested concept (Miles, 2017) – which creates difficulties for theoretical and empirical research. The review by Miles (2017) stresses that stakeholders’ boundaries

and stakeholder identification are essential, and proposes four stakeholder classes: influencer, claimant, recipient, and collaborator. This classification could be adapted to provide a more nuanced understanding of the normative and instrumental influences of stakeholders in the context of digitalisation.

Another concept applied by Schons and Steinmeier (2016), for example, is stakeholder proximity, which describes the spatial nearness of stakeholders to a firm, thus highlighting the important role of employees. Stakeholder proximity is characterised by the level of involvement in companies' processes. Differentiating between high- and low-proximity stakeholders can be vital for identifying and applying symbolic versus substantive actions and practices (Schons & Steinmeier, 2016). Finally, future research may explicitly apply constructs like stakeholder engagement, which 'refers to the aims, activities, and impacts of stakeholder relations in a moral, strategic, and/or pragmatic manner' (see Kujala et al., 2022 for a review). Measuring stakeholder engagement is important to underscore its theoretical development, however, as Kujala et al. (2022) outlined, most studies of stakeholder engagement adopt qualitative methods, whereas they believe that applying quantitative models would be more appropriate to map the different relationships across and within stakeholder groups. Hence Kujala et al. (2022) call for refocusing on the relational view of stakeholder engagement instead of the transactional and entity-focused process, and to account for potential interdependencies between stakeholders, as involving some groups may affect engagement with others. An integrative view of stakeholder engagement also includes addressing the 'dark side' of stakeholder engagement and its intended and unintended consequences, such as conflicting views, misalignments, or false claims (Kujala et al., 2022). This is an important topic for CDR and CSR research: to identify where unavoidable conflicts arise and how to resolve them.

Stakeholder theory, or constructs such as stakeholder engagement, may advance the field and lead to a broader research perspective by incorporating all stakeholder groups, revealing promising research questions. For instance, what are the consequences, changes, and expectations within a supply chain after digitalisation? How can an ecosystem be designed in a participatory way to align with the circular economy? How is CDR/ CSR perceived by and applied to employees working remotely? This also calls for diversifying methods and using qualitative and quantitative research constructs, measurements, and approaches.

Figure 9: Potential Research Paths

Key Findings or Challenges	Examples of Research Questions
Fragmented Perspectives	**How can the two concepts be reconciled?** • How can the different perspectives (CDR and CSR) be integrated in a holistic corporate strategy? How to promote cross-disciplinary approaches? • Are there any interdependencies between different dimensions and how are they related • How can trade-offs be managed in practice?
Artificial Actors	**Who takes responsibility for corporate actions in the digitised world?** • How do companies understand their responsibility, and how do/can they identify the AI element in decision-making? • How are/can social and ecological criteria be introduced into AI-supported decision-making? • Who influences and decides upon the criteria for algorithms?
Data Privacy & Digital Trust	**How can responsible digitalisation be institutionalised in corporate actions and policies? What are the expectations of stakeholders?** • How do (digital) trust and transparency interrelate? • Is the term 'corporate' social or digital responsibility still appropriate for the platform economy? • (How) do the social and digital responsibilities of companies supplying and applying digital technologies differ?
Stakeholder Theory	**What do we learn from CDR and CSR in digitalisation if applying stakeholder theory and a stakeholder perspective?** • How can businesses/stakeholders build a relationship with non-human actors? How to explore stakeholder engagement as a practice in different digital settings. • How does the relationship with diverse stakeholder differ depending on the company's digital maturity level? • (How) does the stakeholder-proximity change when moving from a non-digital to a digital context, and between purely digital stakeholders, purely analogous ones, and those combining both? How does this affect CDR/CSR, symbolically and substantively?
Self-Regulation	**How can corporations genuinely implement CDR and CSR in parallel?** • How can self-commitment be implemented? • How to avoid machinewashing or ethical washing? How can companies walk the talk – providing symbolic and substantial voluntary self-commitment? • How can evolving regulations and self-regulation in the digital context interrelate, and how do they shape CSR/CDR?
Method	**What measurements can be applied to digital technologies and digitalisation?** • What multi-dimensional approaches can be applied, incorporating environmental, economic, social, and digital dimensions? • How do we develop validated quantitative measures of digitalisation and CDR? • How can both CDR and CSR be better integrated and harvested for their interdisciplinary insights?

Taking a Validated and Nuanced Approach to the Role of Digitalisation

While we observed that the literature on the intersection of digitalisation and CSR is more quantitatively driven than the CDR literature, both fields lack established, validated quantitative measures of digitalisation and CDR. To advance the understanding of these relationships, we propose a stronger integration of both research strands to foster interdisciplinary insights and the development of diverse, validated measures. This would allow for a more systematic assessment of how companies integrate digital technologies into their operations, and whether such integration enhances their responsible behaviour in a digitalised world. Moreover, it is crucial to expand and refine how digitalisation and digital technologies are conceptualised and measured, ensuring a more comprehensive understanding of their implications.

Currently, most studies in this literature examine digitalisation in rather broad terms, acknowledging – but not accounting for – its multifaceted nature. Existing measurement approaches include survey-based self-reported digitalisation in firms (e.g., 'During the past three years our organisation has developed strategic plans to integrate AI tools progressively,' (Ghobakhloo, Asadi, et al., 2023b), index-based proxies reflecting digitalisation in the wider environment (e.g., W. Chen, 2023), and textual analyses such as word frequency counts in corporate reports (e.g., Nie et al., 2024). To understand which digital technologies are discussed, we applied natural language processing (NLP) techniques in R to systematically analyse the identified papers, normalising digital technology mentions relative to total word counts, and employed an extended glossary derived from (B. Li et al., 2024) and enhanced by synonym expansion using OpenAI's ChatGPT-4-turbo (for the glossary, see Appendix A5, p. 220). The results highlight that AI (2.51%) and blockchain (0.8%) are the most frequently mentioned technologies, while digital technology applications (0.36%), big data (0.05%), and cloud computing (0.05%) appear less frequently in the analysed literature. Thus, compared to the quantitative measurements, the theoretical discussions are more diverse, but at the same also still focused on AI.

Considering the current literature, one crucial, yet underexplored aspect is that digitalisation and digital technologies may not have uniform effects on how responsible companies act; rather, they could have varying – even countervailing – driver, outcomes and attributing effects. For instance, while AI can enhance transparency and efficiency in

sustainability reporting, it may also introduce ethical concerns regarding bias and job displacement. Blockchain, often associated with enhanced accountability through immutable records, might simultaneously enable regulatory circumvention. Furthermore, measuring whether and which policies or guiding principles prevail in companies does not indicate whether companies have also implemented corresponding practices, i.e., whether they are walking the talk. For future research, it would be worthwhile to explicitly distinguish between CDR/CSR policies and actual corporate practices, as this is currently rarely done, and which remains a critical gap in the literature. Hence, we propose the question of how we can distinguish between companies that truly integrate digital technologies for responsible innovation and those that merely engage in symbolic adoption? Future research should move beyond surface-level measurements of digitalisation and explore whether companies genuinely 'walk the talk' by implementing meaningful, technology-driven responsibility practices. This calls for novel methodologies that assess both the intended and unintended consequences of digital transformation for corporate responsibility.

3.6 Conclusion

This study synthesises and compares the research topics and approaches towards CDR and CSR by employing a systematic literature review, which identifies key differences and commonalities. The analysis has shown how scholars are approaching the challenges of businesses responsibly navigating the changes induced by digital transformation, and their implications for the conceptualization and study of CDR and CSR. Our findings reveal that CDR research is highly interdisciplinary, combining perspectives from information systems and business ethics (among others), which allows adds depth to the normative discussions of the issues and challenges. The current CSR literature, by contrast, mostly comprises empirical investigations into the interrelations and performance aspects of CSR and digitalisation at the management level. While both concepts address specific topics, they also share important areas of common ground. By translating CDR and CSR into a unified corporate strategy, businesses can align the responsibilities, challenges, and opportunities arising from the interrelation between digital technologies and sustainable transformation.

CHAPTER 4 | Peers: Powerful or Negligible? A Systematic Review on Peer Factors and Internal Whistleblowing

Behnud Mir Djawadi Sabrina Plaß Sabrina Loer

When reporting wrongdoing internally, whistleblowers are confronted with the dilemma of weighing up their loyalty towards the organization (e.g., ethical standards) and their co-workers (e.g., the social norm of not snitching on peers). However, the role played by peers in the whistleblowing decision process and in the aftermath has rarely been addressed in existing reviews. We therefore perform a systematic review that identifies seven thematic clusters of peer factors, offering researchers an informative overview of (a) the peer factors that have been examined to influence the whistleblowing decision, and (b) the extent to which the whistleblower experiences adverse consequences from peers in the aftermath of whistleblowing. As peer factors seem to be important to explain and predict internal whistleblowing, researchers are encouraged to address in future works the research gaps our review unravelled.

Keywords: *Whistleblowing, Peer Factors, Co-workers, Systematic Review*

4.1 Introduction

One way of mitigating wrongdoing in organizations is to encourage employees to blow the whistle on observed unethical behaviour (Keenan, 2000). We define whistleblowing as "the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action" (Near & Miceli, 1985:4). Regulators and managers have increasingly become aware of the benefits of wrongdoing being reported internally before it is publicly exposed. In contrast to reporting observed wrongdoing through external channels, internal whistleblowing enables organizations to address and correct wrongdoing themselves and minimizes reputational damage (Lee & Xiao, 2018). Whistleblowing has thus become an accepted part of the regulatory environment of organizations, to ensure legal compliance and ethical business practice (Vandekerckhove, Brown, & Tsahuridu, 2014). As such, organizations increasingly adopt and establish whistleblowing programs as part of their formal compliance systems, which include e.g., policies and reporting channels to structure and facilitate the reporting process and make the organization's commitment to whistleblowing more transparent (Dixon, 2016; Hassink et al., 2007).

However, to engage in internal whistleblowing, employees must be willing to report on organizational members with whom they stand in a direct or indirect professional and/or personal relationship (Trevino & Victor, 1992). Especially given the growing trend whereby organizations invest in team building and set up working groups (Oh et al., 2004), the potential whistleblower would have to consider turning on colleagues, which would cause a dilemma and threaten the strong social ties that organizations try so hard to foster. Hence, whistleblowing scholarship has seen a shift in the way internal whistleblowing is framed: it is no longer merely an act of ethical resistance aimed at changing corporate or governmental behaviour and, in the process, being met with crude antagonism by the organization (Glazer & Glazer, 1989). Rather, whistleblowing is recognized more as a complex social phenomenon, due to whistleblowers having to weigh up their loyalty towards the organization and its ethical standards against their relationship with co-workers (e.g., by conforming to the prevailing social norm by not snitching on peers). In this regard, the role of peers in the process and

in the aftermath of the whistleblowing decision has more recently become a non-trivial dimension in whistleblowing research (S. R. Khan & Howe, 2021; Mayer et al., 2013a).

Moreover, studies based on different theories, such as social identity theory (Anvari et al., 2019) and social information processing (Gundlach et al., 2003; Near & Miceli, 1995), stress the importance of the immediate workgroup for a whistleblower to assess the observed situation as wrongdoing, and its reporting as a reasonable action. Due to their direct proximity, peers are particularly salient and satisfy individuals' social needs (Greenberger et al., 1987), and hence may have a formative impact on a whistleblower's decision. McLain and Keenan (1999) even pose the theoretical argument that in a perceived conflict over whistleblowing between the organization's response and co-workers' responses, employees are more likely to behave in a way that pleases the peers rather than the organization. Muehlheusser and Roider (2008) similarly propose that peers place social loyalty over integrity, resulting in a wall of silence, where peers will not be reported, despite an organization's policy promoting the reporting of misconduct. As these theoretical studies demonstrate, over-prioritizing the relationship with peers in one's immediate working environment might hamper the attempt to blow the whistle on observed wrongdoing (Dungan et al., 2019).

Therefore, this handbook chapter asks how different peer factors are related to internal whistleblowing. We hereby complement existing reviews (Culiberg & Mihelič, 2017; Mesmer-Magnus & Viswesvaran, 2005), which have focused on individual, situational, and organizational factors of whistleblowing. Performing a systematic literature review allows us to identify which peer factors have already been investigated and which have potential for further research (Rousseau et al., 2008; Tranfield et al., 2003). We consider this method appropriate because the role of peers in the context of whistleblowing has already been investigated in studies based on different constructs and with specific foci. However, the studies are fragmented and disconnected, notably because they evolved in different contexts, such as academia, accounting, business, the military, or nursing.

Our review makes several contributions. First, we provide a framework that structures and categorizes the existing literature on peer factors and internal whistleblowing into two main strands: (a) the peer factors that have been examined to influence the whistleblowing decision, and (b) the extent to which the whistleblower

experiences adverse consequences from peers in the aftermath. While the effect of peers has often been simplified in one variable, such as co-worker support, we find that peer factors are far more multifaceted. Our framework groups peer factors thematically into seven further clusters, offering researchers an informative overview of the factors studied within each category along with their (mixed) empirical findings. Second, we suggest that peer factors interact with other types of variables (i.e., moderating effects), so that specific relationships might only persist under particular social conditions. Third, we identify that further research is needed in order to unravel how and why peer factors influence the whistleblowing decision, and whether peer factors can explain relationships between previously investigated variables and whistleblowing (i.e., mediating effects). Lastly, by better understanding the different patterns of negative consequences for whistleblowers, researchers will be able to differentiate formal organizational variables (e.g., work-related retaliation) from informal peer variables (e.g., peer ostracism) which, in turn, could help managers develop more fine-grained programs for the prevention of these consequences and the promotion of whistleblowing.

4.2. Shortcomings of Peer Factor Studies in Previous Whistleblowing Literature Reviews

Similar to Near and Miceli (1995), we refer to peers as co-workers or colleagues in the immediate workgroup with whom potential whistleblowers share at least a professional relationship. Peer factors are by and large intangible, hard to observe, and only indirectly controllable by the organization. They refer to, for example, the informal social norms, implicit agreements, and informal relationships. Peer factors, then, form the set of variables in the immediate relational and social context of potential whistleblowers, and can be assigned to the informal social structure of the organization, i.e., to each respective working group (Murphy, 2021).

Although theoretical work suggests that peer factors serve as valid antecedents for whistleblowing and affect the whistleblower in the aftermath, peer factors as set of variables have rarely been addressed in existing whistleblowing reviews. Over a decade ago, a meta-analysis of whistleblowing factors conducted by Mesmer-Magnus and Viswesvaran (2005) provided an essential momentum for whistleblowing research[27].

[27] Whistleblowing research includes a vast array of studies. Different theoretical models address the question of the stages in the decision-making process that potential whistleblowers go through to decide

Their review examined 26 empirical studies identifying the personal, situational, and contextual predictors for the internal and external whistleblowing intention, for actual whistleblowing behaviour, and retaliation. However, the authors mainly treat the influence by peers as a composite variable for organizational retaliation. We argue that responses by peers differ from work-related retaliation. While the former manifest in the form of informal social sanctions (e.g., ostracism), mostly by co-workers (Williams, 1997, 2001), work-related retaliation takes the form of poor performance reviews, relocation, or suspension from the job. Yet, as Hollinger and Clark (1982) show, employees are far more susceptible to – and more likely to be deterred by – the fear of social threats and sanctions emanating from co-workers, than by the organization's formal sanctions.

Likewise, subsequent literature reviews that built on this seminal meta-analysis did not focus on peer factors either. Vadera et al. (2009), for example, reveal that situational and contextual determinants predict whistleblowing outcomes more consistently than personality factors. Culiberg and Mihelič (2017) identify further research gaps by developing a framework that captures a comprehensive set of variables concerning who the whistleblower is, how, and why they report the wrongdoing, and to whom. Gao and Brink (2017) use the model proposed by Near and Miceli (1995) to review whistleblowing studies in the context of accounting-related misconduct. The authors cluster the determinants of whistleblowing intentions, reviewing the relevant empirical findings for each determinant. Lee and Xiao (2018) expand the scope of determinants (e.g., whistleblowing legislation) in the context of accounting-related misconduct and show that predictors for internal and external whistleblowing intentions differ.

Variables on the group level, such as group cohesiveness that cannot be assigned to either organizational or individual factors, have not been addressed in any of these reviews. While some reviews (Vadera et al., 2009) partly examine the influence of the

whether to report wrongdoing (Alleyne et al., 2013; Dozier & Miceli, 1985; McLain & Keenan, 1999; Miceli & Near, 1992). Other studies examine how individual characteristics such as demographic features or moral personality traits distinguish potential whistleblowers from silent observers (Liyanarachchi & Newdick, 2009a). Researchers investigate how the characteristics of the wrongdoing (i.e., type and seriousness), of the wrongdoer, or the fear of retaliation affect the whistleblowing decision (Cassematis & Wortley, 2013), often referred to as situational factors. Studies on organizational factors examine how the characteristics of organizations (e.g., organizational climate), (ethical) leadership and changes in organizations' policies to protect whistleblowers and facilitate the reporting process can either increase or hamper the willingness of whistleblowing (Kaptein, 2011).

organizational climate – by exploring employees' perception of the organization's ethical standards – they do not capture the attitude of peers. This aggregated categorization may stem from the assumption that the ethos underpinning the organizational response to whistleblowing is shared equally among the whistleblower's peers. As such, variables pertaining to the whole organization are assumed to be conceptually equivalent to variables on the group level. However, as discussed in several studies (Tenbrunsel et al., 2003), an organizational climate can fundamentally differ from the norms developed and established informally in a working group. To the best of our knowledge, the only review that explicitly mentions peer factors (as 'social factors') as further determinants is that by Nicholls et al. (2021), who connect different literature strands (e.g., in an accounting and a non-accounting context) and extend the group of potential determinants of whistleblowing intentions. However, because their review comprises seven other sets of variables related to the whistleblowing intention (e.g., cost-benefit motives, expectation of whistleblowing consequences), the number of identified peer factors is quite limited (mainly, norms, group structure, and support for the whistleblower). Furthermore, their review includes studies that refer to the whole organization rather than to workgroups of peers (e.g., studies about whistleblower support).

In contrast to the published reviews, our systematic review focuses on variables on the peer level. Hence, we refer to whistleblowing as a social phenomenon, with the whistleblower firmly located in a network of non-trivial relationships, in which they consider the consequences of their decision for various parties, and weigh up the often competing interests such as those of the organization, superiors, or peers (S. R. Khan & Howe, 2021). With our review, we seek to complement the research on organizational, situational, and individual determinants to advance the understanding of the factors that enhance or hinder internal whistleblowing.

4.3 Method

We guided our systematic literature review along the three main stages exemplified by Tranfield et al. (2003), and Thorpe et al. (2005), namely, review planning, conducting, and reporting and dissemination of findings.

Using Scopus, Web of Science, and EBSCOhost as a database we identified a total of 788 articles for the keywords whistle* AND peer* OR cowork* OR co-work* OR team* OR colleague. In a first screening, we removed 200 duplicates, before performing

several rounds with exclusion criteria excluding further in total 479 studies. We only proceeded with peer-reviewed journal articles that used primary empirical research (i.e., excluding handbooks). After a detailed review, we excluded further 83 studies that did not fit the peer variable or the underlying whistleblowing definition as defined for this review. Note that we count peer reporting as a sub-category of whistleblowing following the conceptualization of peer reporting as a type of whistleblowing by Trevino and Victor (1992). Additionally, we included seven articles that fit the criteria and that were mentioned as a reference in the reviewed papers. This resulted in a total of 33 studies.

4.4 Review Results

To analyse the 33 retained studies, we identified two overarching categories under which we clustered the results, namely, whether peer factors are described as an influence over (=antecedent) or as a consequence of (=aftermath) the whistleblowing decision. Within these overarching categories, we assigned all identified factors to a second level of thematic categories (see Figure 1), but only for structuring purposes, as they do not represent a theoretically-based categorisation, and are overlapping.

Peer Influences on Whistleblowing

The study characteristics and findings on peer influences are summarized in Table 1. In total, this category comprises 26 studies conducted in diverse countries and cultures (e.g., western and Asian). Moreover, the underlying contexts and samples vary greatly, including, for example, athletes, military personnel, auditors, students, and employees. Sample sizes range between 15 and 10,850. Twenty-two articles used quantitative methods, three employed qualitative methods, and one article used both. Half of the studies examined the direct relationship between the corresponding peer variable and whistleblowing. Of these, the majority (17) investigate whistleblowing *intentions*, seven analyse actual whistleblowing *behaviour*, and two consider both intentions and actual behaviour.

In the context of peer influences, we found five categories to which we assigned the identified peer factors: (I) peer involvement in wrongdoing and whistleblowing situations, (II) allegiance to peers and the organization, (III) behavioural prescription by peers, (IV) relationship and experiences with peers, and (V) fear of consequences from peers (see Figure 1).

Figure 1: Framework with Seven Sub-Categories of Peer Factors and Internal Whistleblowing (Influences & Consequences)

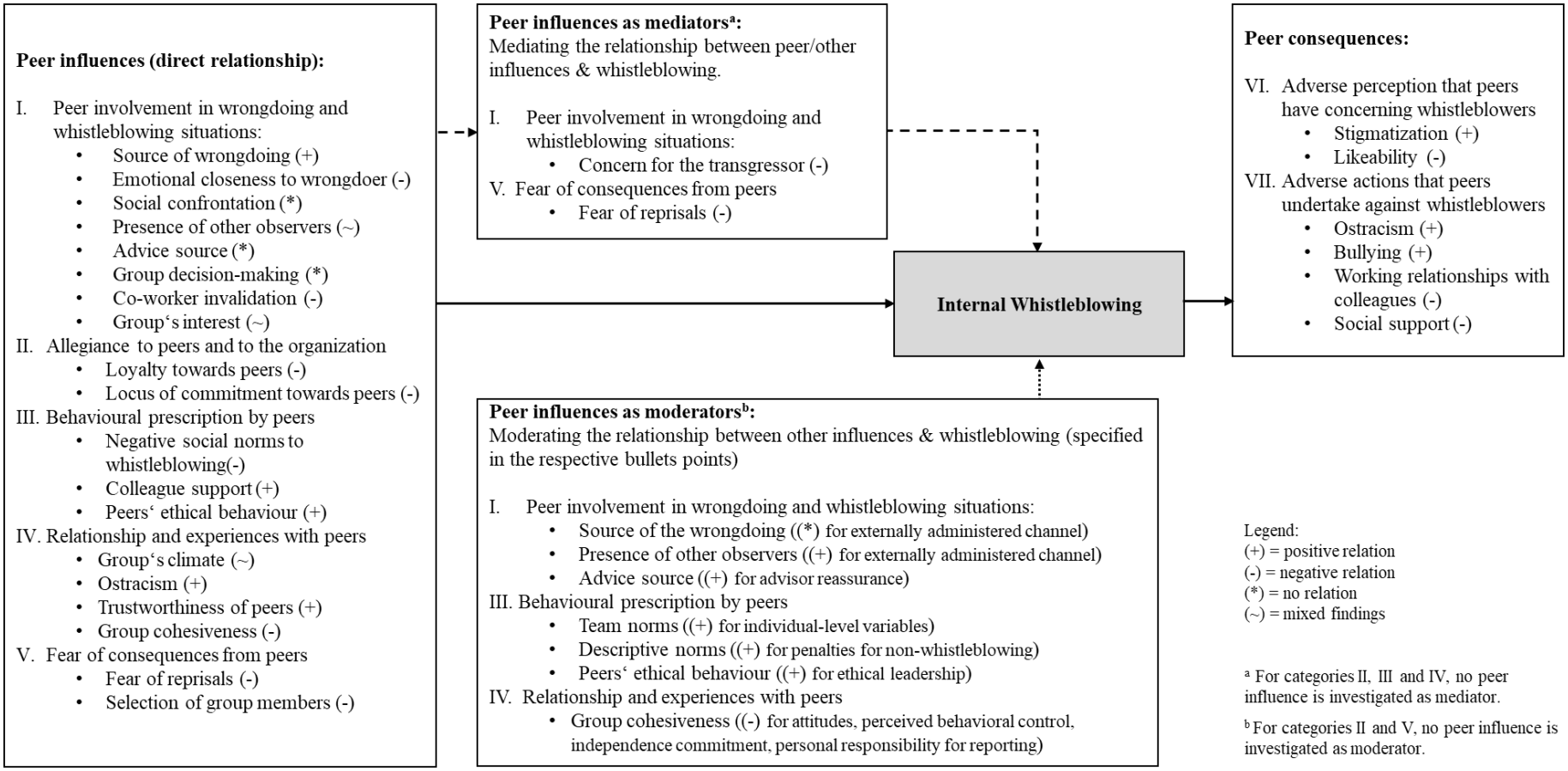


Table 1: Identified Articles about Peer Influences on Internal Whistleblowing

Author(s) (year)	Context	Country	Sample (size)	Method	Peer factor(s)	Category	Direct relation-ship / mediator/ moderator	Intention or beha-viour	Main results
Afe, et al. (2019)	Wrongdoing in organizations	Turkey	Academic personnel (250)	Question-naire (quant.)	Mobbing	Relationship and experiences with peers	Direct	Intention	Persons who perceive mobbing by peers are more willing to blow the whistle on their peers' unethical act informally (no effect for formal whistleblowing)
Alleyne, et al. (2019)	Fraud in organizations	Barbados	Auditors (226)	Survey, scenario (quant.)	Group cohesiveness	Relationship and experiences with peers	Moderator	Intention	Strong group cohesion negatively interacts with predictors of whistleblowing
Barkou-kis, et al. (2021)	Doping in sports	Cyprus	Sport athletes, coaches, sport directors (15)	Structured interviews (qual.)	Perceived social norms towards whistle-blowing	Behavioural prescription by peers	Direct	Intention	Perceived social norms towards whistleblowing as a deterrent for whistleblowing: Especially in small community, people attribute negative connotations to whistleblowing
Boo, et al. (2021)	Fraud in organizations	Singa-pore	Senior auditors (69)	Scenario experiment, question-naire (quant.)	Advice source	Peer involvement in wrongdoing and whistleblowing situations	Direct, moderator	Intention	Whistleblowing intention does not differ depending on whether advice on the whistleblowing situation comes from the technical department as an authoritative source or from a colleague as a non-authoritative source No significant interaction with advisor reassurance
Chang, et al. (2017)	Corruption in organizations	South Korea	Government employees (5,706)	Survey (quant.)	Colleague support	Behavioural prescription by peers	Direct	Intention	Positive relationship between perceptions about colleague support and whistleblowing intentions
Chen, et al. (2017)	/	/	Students (147)	Laboratory experiment (quant.)	Descriptive norms	Behavioural prescription by peers	Moderator	Behaviour	Interaction between incentive framing and descriptive norms: penalties lead to a greater increase in whistleblowing (compared to rewards) when descriptive norms supporting whistleblowing are

									strong than when descriptive norms supporting whistleblowing are weak
Curphy, et al. (1998)	Violation of honour code	USA	Military employees (365)	Scenarios (quant.)	Emotional closeness to wrongdoer; Presence of other witnesses	Peer involvement in wrongdoing and whistleblowing situations	Emotional closeness to wrongdoer: direct; Presence of other witnesses: direct	Intention	Emotional closeness to wrongdoer: wrongdoing committed by peers who are close friends is less likely reported than wrongdoing committed by peers who are relative strangers Presence of other witnesses: higher reporting intentions when others also observed the wrongdoing
Gao, et al. (2015)	Fraud in organizations	USA	Students (369)	Scenario experiment (quant.)	Bystander effect (presence of bystanders); Wrongdoer (superior vs. co-worker)	Peer involvement in wrongdoing and whistleblowing situations	Bystander effect: direct, moderator; Wrongdoer (superior vs. co-worker): direct, moderator	Intention	Bystander effect that negatively impacts whistleblowing intentions in one scenario and a positive impact in the other scenario; An individual's whistleblowing intention is significantly lower when the wrongdoer is a supervisor than when s/he is a co-worker; The source of wrongdoing does not significantly influence the positive relationship between an externally administered reporting channel and whistleblowing; The presence of other bystanders positively influences the positive relationship between an externally administered reporting channel and whistleblowing
Goddik-sen, et al. (2021)	Academic cheating	Denmark Hungary Ireland	Students (72)	Qualitative interviews, scenarios (qual.)	Loyalty towards peers; Fear of negative responses from peers	Loyalty towards peers: Allegiance to peers and the organization; Fear of negative responses from peers: Fear of consequences from peers	Direct	Intention	Loyalty leads to direct, personal confrontation with wrongdoer (not whistleblowing) Negative reactions from peers

Iwai, et al. (2021)	Academic cheating	Brazil	Students (947)	Question-naire, scenarios (quant.)	Peer ethical behaviour; Fer of reprisal from one's peers	Peer ethical behaviour: Behavioural prescription by peers; Fear of reprisal from one's peers: Fear of consequences from peers	Peer ethical behaviour: direct; Fear of reprisal from one's peers: mediator	Intention	Effects of peer ethical behaviour on whistleblowing intentions are mediated by fear of retaliation. Fear of retaliation from peers as mediator: peer ethical behaviour has negative effect on fear of retaliation
Kaplan, et al. (2010)	Fraud in organizations	/	Students (96)	Scenario experiment (quant.)	Unsuccessful social confrontation	Peer involvement in wrongdoing and whistleblowing situations	Direct	Intention	After unsuccessful meeting with the transgressor, intentions to report the wrongdoing to the wrongdoer's supervisor are significantly higher than reporting intentions to an internal auditor; reporting intentions do not differ concerning the types of recipients without social confrontation; unsuccessful social confrontation does not generally affect reporting intentions
Khan & Howe (2021)	Fraud in organizations	USA	Study 1: students; (187) Study 2: MTurk participants (375)	Scenario experiment (quant.)	Group cohesiveness ; Concern for transgressor	Group cohesiveness: Relationship and experiences with peers; Concern for transgressor: Peer involvement in wrongdoing and whistleblowing situations	Group cohesive-ness: direct; Concern for trans-gressor: mediator	Intention	High group cohesiveness increases concern for the wrongdoer and consequently reduces the likelihood of whistleblowing Increased concern for the wrongdoer decreased the likelihood of reporting
Latan, et al. (2018)	Fraud in organizations	Indonesia	Public accountants (256)	Question-naire, scenarios (quant.)	Team norms	Behavioural prescription by peers	Moderator	Intention	Team norms partially moderate relationship of individual-level variables (attitudes toward whistleblowing, perceived behavioural control, independence commitment, personal responsibility for reporting, and personal cost of reporting) with

									(internal and external) whistleblowing intentions
Mayer, et al. (2013)	Unethical behaviour in organizations	USA	Employees (197)	Question-naire, field study, experiment (quant.)	Co-workers' ethical behaviour	Behavioural prescription by peers	Moderator	Intention / Behaviour	Interaction between supervisory ethical leadership and co-worker ethical behaviour on internal whistleblowing
McIntosh, et al. (2019)	Wrongdoing in organizations	USA	Students (534)	Scenario (quant.)	Source of unethical behaviour (peer or advisor)	Peer involvement in wrongdoing and whistleblowing situations	Direct	Intention	Participants less likely to report the observed misconduct of an advisor compared to a peer
Miceli, et al. (1991)	Academic misconduct	USA	Students (295)	Field experiment (quant.)	Number of observers of wrongdoing	Peer involvement in wrongdoing and whistleblowing situations	Direct	Behaviour	Observers of wrongdoing more likely to blow the whistle when more, rather than fewer, other observers were present
Miceli, et al. (2012)	Wrongdoing in organizations	USA	Military and civilian employees (3288)	Question-naire (quant.)	Co-worker invalidation	Peer involvement in wrongdoing and whistleblowing situations	Direct	Behaviour	Less perceived invalidation of whistleblowing by co-workers predicted whistleblowing
Pershing (2002)	Violation of honour code	USA	Military employees (527; 40)	Survey, semi-structured interviews (qual.)	Loyalty	Allegiance to peers and the organization	Direct	Behaviour	Peer loyalty results into non-reporting for occupational misconduct; “code of silence” impedes direct reporting even with an Honour Concept in place, and forms of counselling the perpetrators are chosen to uphold both loyalties
Rennie & Crosby (2002)	Academic misconduct	Scotland	Students (676)	Question-naire; focus groups	Fear of negative peer reactions	Fear of consequences from peers	Direct	Intention	Fear of negative reactions from peers negatively relate to whistleblowing
Reuben & Stephen-son (2013)	/	/	(68)	Laboratory experiment (quant.)	(Anticipation of) group in- & exclusion (selection)	Fear of consequences from peers	Direct	Behaviour	Option that peers can select who is included in a group for future cooperation reduces probability of reporting wrongdoing
Rothwell & Baldwin	Violations of rules	USA	Police officers; civilian	Survey, vignettes (quant.)	Friendship or team climate	Relationship and experiences with peers	Direct	Intention / Behaviour	Team climate positively related to whistleblowing intentions but unrelated to whistleblowing behaviour

(2007)			employees (382)						
Spoelma, et al. (2021)	Unethical behaviour in organizations	USA	Employees (109); Students (108)	Surveys, experiment (quant.)	Ostracism	Relationship and experiences with peers	Direct	Behaviour	Ostracism has a positive effect on whistleblowing
Taylor (2018)	Wrongdoing in organizations	Australia	Employees (10,850)	Australian government data (quant.)	Trustworthiness of co-workers	Relationship and experiences with peers	Direct	Behaviour	Internal whistleblowing is positively related to perceptions of trustworthy co-workers
Taylor & Curtis (2010)	Fraud in organizations	/	Senior auditors (120)	Vignette scenarios (quant.)	Locus of Commit-ment; organization vs. co-workers	Allegiance to peers and the organization	Direct	Intention	As an individual's commitment moves toward the organization and away from colleagues, likelihood of reporting and perseverance increase
Taylor & Curtis (2013)	Fraud in organizations	USA	Senior auditors (106)	Vignette (quant.)	Reporting peer vs. supervisor	Peer involvement in wrongdoing and whistleblowing	Direct	Intention	Positive relationship to whistleblowing intention when wrongdoer is a peer rather than a supervisor
Trevino & Victor (1992)	Academic cheating; theft at workplace	USA	Students (478; 115); Employee (128)	Scenarios, field study (quant.)	Group members' interests	Peer involvement in wrongdoing a whistleblowing	Direct	Intention	If group members are negatively affected by the wrongdoing, the inclination to blow the whistle increases

Peer Involvement in Wrongdoing and Whistleblowing Situations

The first category, *peer involvement in wrongdoing and whistleblowing situations,* synthesizes studies investigating factors that relate to peers directly involved in either the wrongdoing or the whistleblowing situation.

Three studies on peer involvement in wrongdoing situations focused on whether the *source of the wrongdoing* influences the whistleblowing decision (Gao et al., 2015; McIntosh et al., 2019; Taylor & Curtis, 2013), that is, whether the wrongdoing is committed by a peer or by a superior. All three studies consistently find that intentions for whistleblowing are more likely if the perpetrator is a peer rather than a supervisor. While reporting a peer might be perceived as less of a threat, the anticipated consequences of reporting an influential advisor tend to be higher (McIntosh et al., 2019). Examined as moderator, the source of wrongdoing does not significantly influence the positive relationship between an externally administered reporting channel and whistleblowing (Gao et al., 2015).

Among peers who commit the wrongdoing, Curphy et al. (1998) investigate the relationship between *emotional closeness to the wrongdoer* and the whistleblowing intention with a sample of cadets from the U.S. Air Force Academy. They find that, if the wrongdoers are close friends, they are less likely to be reported than if they are relative strangers, even though cadets are expected to follow the Academy's honour code (a formalized norm for reporting) (Curphy et al., 1998).

Kaplan et al. (2010) set up an experiment with students to examine how an unsuccessful *social confrontation* with the wrongdoer influences to whom a potential whistleblower will report, finding that, after an unsuccessful meeting with the transgressor to discuss the wrongdoing, the whistleblower is more likely to report to the wrongdoer's supervisor than to an internal auditor. Drawing on power theories, the authors state that an unsuccessful confrontation increases the inclination to report to a powerful recipient (i.e., the wrongdoer's supervisor). By contrast, when a social confrontation does not take place, reporting intentions do not differ in terms of the choice of the recipient. Moreover, unsuccessful social confrontation does not generally affect reporting intentions.

Lastly, Khan and Howe (2021) elaborate on how a whistleblower is affected by their *concern for the transgressor (wrongdoer)*. In their experiment with students,

participants who are asked to imagine a wrongdoing committed by a close group member report significantly greater concern for the transgressor, and, more specifically, for their potential suffering as a consequence of the whistleblowing. This increased concern for the transgressor decreases the likelihood of reporting them. Thus, the concern for the transgressor is identified as a mediating variable of the relationship between group cohesiveness and whistleblowing.

In the following, we consider studies where peers may influence the whistleblowing decision by their sheer presence or by affecting the potential whistleblower's decision-making process. Three studies address the *presence of other observers* of the wrongdoing (Curphy et al., 1998; Gao et al., 2015; Miceli et al., 1991). The results seem inconsistent. Curphy et al. (1998) find that reporting intentions are more likely when the wrongdoing is also observed by others. Miceli et al. (1991) further indicate that whistleblowing increases with the number of observers. By contrast, in one of their scenarios, Gao et al. (2015) provide indication for the bystander effect: when more than one person observes the wrongdoing, whistleblowing by any of the individual observers becomes less likely due to the diffusion of responsibility. The bystander effect is only apparent, however, if the reporting channel is administered internally, rather than externally. The presence of other bystanders enhances the positive relationship between an externally administered reporting channel and whistleblowing. In another scenario, the same authors (Gao et al., 2015) find that the presence of another observer impacts whistleblowing intentions positively.

O'Leary and Pangemanan (2007) analyse *whether individual or group decision-making* leads to ethical behaviour in the form of whistleblowing. The study finds that groups are more likely to come to a neutral decision, i.e., ignoring an observed wrongdoing, whereas individuals make more extreme decisions either in the unethical direction of participating in the wrongdoing or in the ethical direction of reporting it.

Miceli et al. (2012) investigate, among others, the influence on whistleblowing exerted by a potential whistleblower's perception of how their co-workers perceive the wrongdoing and whether it should be reported. Their results reveal that the perceived *co-worker invalidation* is negatively associated with whistleblowing. This provides empirical evidence for the crucial influence of co-workers on whistleblowers, and on how whistleblowers justify their decision to report, or not. In three separate studies, Trevino

and Victor (1992) examine the situation when the *group's interests* are negatively affected by the wrongdoing. In two of their studies (one in an academic context and the other in a fast-food restaurant), they observe that the inclination to blow the whistle increases if group members are negatively affected by the wrongdoing. They find no support for this relationship in their third study (also a fast-food restaurant context).

Finally, there is no support for whether the type of *advice source* is of particular importance. In an experiment with participants from the "Big 4" firms (e.g., KPMG), Boo et al. (2021) show that the whistleblowing intention does not depend on whether advice on the whistleblowing situation comes from the technical department, as an authoritative source, or from a colleague, as a non-authoritative source. Further, the interaction between advice source and advisory reassurance is statistically insignificant.

To summarize, the evidence in the category *peer involvement in wrongdoing and whistleblowing situation* is mixed: while the reviewed studies provide support for reporting a peer rather than a superior, the impact of the presence of observers varies. A potential reason for the inconsistency of these findings might be the different study settings and designs. While Curphy et al. (1998) examined these effects in a military institution, where an honour code guides cadets' behaviours, both Miceli et al. (1991) and Gao et al. (2015) ran a student experiment. The validity of the findings from these studies, e.g., *emotional closeness, concern for the transgressor* or *social confrontation,* would benefit from being replicated.

Allegiance to Peers and to the Organization

The second category, *allegiance to peers and to the organization*, spans studies that investigate factors relating to potential whistleblowers' allegiance to peers and/or to their organization.

Goddiksen at al. (2021) identify that *loyalty* considerations towards peers are an important reason for their reluctance to report. However, rather than leading to inaction, loyalty considerations are more likely to lead to a direct, personal confrontation with the wrongdoer. The influence of peer loyalty was also highlighted by Pershing (2002), who linked the non-reporting of occupational misconduct to a "code of silence", which forms a central part of the culture of the Naval Academy, and impedes the reporting even when an Honour Concept is in place. Confronting the perpetrator in private is the preferred option as it upholds loyalty both to the organization and to the peer.

Similar to the concept of loyalty, Taylor and Curtis (2010) deal with the relationship between *locus of commitment* and whistleblowing. In this context, the authors distinguish between commitment to the organization and to colleagues. Organizational commitment refers to the "strength of employees' identification with and involvement in a particular organization, a strong belief in organizational goals and values, and a willingness to exert considerable effort on behalf of the organization" (Taylor & Curtis, 2010: 24). Colleague commitment involves "a sense of responsibility, reliability, and readiness to support colleagues within an organization" (Taylor & Curtis, 2010: 24). They find that, the stronger the employee is committed to the organization – rather than to colleagues – the greater the likelihood of whistleblowing, hence linking commitment to colleagues negatively to the reporting of wrongdoing.

It becomes apparent, then, that all three outlined studies support the notion that peer loyalty and commitment inhibit whistleblowing, and, the stronger one's allegiance to peers, as opposed to the organization, the less likely whistleblowing becomes.

Behavioural Prescription by Peers

In the category *behavioural prescription by peers*, we summarize studies dealing with factors of how (perceived) expectations and behaviours by peers influence a potential whistleblower.

Three studies focus on the role of norms conveyed by peers. Latan et al. (2018) refer to *team norms* as rules that are informally adopted within groups to regulate the behaviour of group members. They find that team norms partially moderate the relationship between several individual-level variables (e.g., attitude toward whistleblowing and perceived behavioural control) with internal and external whistleblowing intentions. Barkoukis et al. (2021) conduct interviews with stakeholders in the sports sector and find that, especially in small communities – where misconduct might be more easily identified and suspected, but reporting perceived as snitching – negative *perceived social norms to whistleblowing* can act as a deterrent for whistleblowing. Chen et al. (2017) address the question of how *descriptive norms* –an individual's perception of the behaviour of others in a certain situation – influence whistleblowing. The authors reveal that descriptive norms interact with incentives for whistleblowing, i.e., when the descriptive norms for whistleblowing are strong rather than weak, sanctions are more effective than rewards in increasing whistleblowing.

Peers' ethical behaviour and *support from colleagues* may likewise influence a whistleblower's perception on whether reporting will be endorsed or disapproved of. Two of the three studies analysed in this regard investigate the relationship between *peers' ethical behaviour* and whistleblowing (Iwai et al., 2019; Mayer et al., 2013). Mayer et al. (2013) operationalize this as the perception of the extent to which peers set an example by following ethical standards and behaviour, and find that peers' ethical behaviour and supervisory ethical leadership positively interact in explaining internal whistleblowing. Iwai et al. (2019) support the positive relationship between peers' ethical behaviour and whistleblowing in an academic context. Another form of signalling whistleblowing as the appropriate behaviour is *colleague support*. In a large survey (n=5706) with governmental employees in South Korea on corruption in organizations, Chang et al. (2017) find a positive relationship between the perceptions that arise from colleagues' positive responses after previous whistleblowing incidents and subsequent whistleblowing intentions.

In summary, the reviewed studies consistently indicate that potential whistleblowers are influenced by their perception of norms conveyed by peers, as well as by their behaviour.

Relationship and Experiences with Peers

The fourth category, *relationship and experiences with peers,* refers to the social ties with peers and perceptions about important similarities and experiences with the other members of the social group.

Two studies examine the influence of the *group's climate* that has been developed through long-term interaction with peers. Rothwell and Baldwin (2007) show in their study with police officers that a team climate, which is expressed by showing concern for the well-being of others, is positively related to whistleblowing *intentions* but unrelated to whistleblowing *behaviour*. A negative team experience is created by, for example, the climate of mobbing in the organization. Afe et al. (2019) examine the impact of such a climate on whistleblowing in an academic context, where they find that students who think they are more likely to be mobbed by peers are more willing to blow the whistle on them. However, this relationship is only observed when peers can blow the whistle informally (e.g., reporting to a close associate who may take further action) rather than formally (e.g., through formal procedures and communication lines for whistleblowing).

Spoelma et al. (2021) focus on the negative experience of *ostracism* – which they define as an individual's perception of being ignored or excluded by peers. Peers that have previously been excluded from the group are more willing to blow the whistle on their former group members than those who still feel that they belong to the group. The whistleblowing motives of the excluded peers tend to be driven by negative reciprocity for perceived bad treatment, rather than by altruistic or moral reasons. By contrast, Taylor (2018) examines how the *trustworthiness of peers*, which she conceptualizes as consisting of the ability, benevolence, and integrity attributed to co-workers, influences the whistleblowing decision. This study reveals that perceptions of trustworthy co-workers are positively linked to internal acts of whistleblowing, but trustworthy senior managers even more so than trustworthy co-workers and supervisors.

Lastly, according to Alleyne et al. (2019), *group cohesiveness* reflects the tendency of group members to form social bonds, creating a sense of group belonging. The authors use group cohesiveness as a moderator variable for the relationships between a set of predictors derived from the theory of planned behaviour (Ajzen, 1991) and whistleblowing intentions on fraud in accounting. In almost all cases, group cohesiveness moderated the relationship between predictor variables and whistleblowing. These relationships prevailed only in groups with low cohesiveness, while they were not apparent in groups with high cohesiveness. Similarly, Khan and Howe (2021) find that high group cohesiveness (group unity) reduces the likelihood of whistleblowing, as long as the wrongdoer is part of the group.

In summary, the described studies show that whistleblowing is influenced by both social ties and by experiences with peers, whether these are positive or negative. Overall, the less attached to the peer group an individual is, the more willing they are to report the wrongdoing.

Fear of Consequences from Peers

The final category, *fear of consequences from peers*, deals with four studies that explicitly examine the potential negative responses of peers towards the whistleblower.

Three studies focus on the *fear of reprisals* or negative responses from peers in general. Goddiksen et al. (2021) Iwai et al. (2019), and Rennie & Crosby (2002) examine this relationship in an academic context, and identify that the fear of retaliation from fellow students negatively relates to the willingness to report on academic fraud. In

addition, the fear of reprisals mediates the relationship between peer ethical behaviour and whistleblowing intentions: even in a social environment where peers are perceived as being committed to ethical values, higher fear of reprisals reduces the willingness to report (Iwai et al., 2021).

In the fourth study, Reuben and Stephenson (2013) examine the relationship between the *selection of group members* and actual whistleblowing behaviour where reporting has real monetary consequences for the peers. In their experimental laboratory study, they find that, when there is an option for peers to select who is included in a group for future cooperation, the probability of reporting wrongdoing decreases. However, as in this study, the potential whistleblower is seen more of a "rat" who reports others for individual gain, this may explain why peers anticipate that group members would not welcome whistleblowers into their group.

In summary, the reviewed studies show the tendency that fear of peer reprisal matters and that these responses may not only affect whistleblowing intentions but actual whistleblowing behaviour.

Peer Consequences after Whistleblowing

Table 2 summarizes the characteristics and findings of studies examining actual peer responses towards the whistleblower in the aftermath of their reporting. The nine identified studies rely on different types of samples (e.g., employees, nurses, or social workers) and sample sizes, ranging from 2 to 83,214 observations. Four of these studies used qualitative, and five used quantitative methods. We divide the factors in these studies into the following two categories: (VI) adverse perception that peers have concerning whistleblowers, comprising stigmatization and likeability, and (VII) adverse actions that peers undertake, such as unofficial reprisals, ostracism, bullying, and the relationship with and social support for whistleblowers (see Figure 1).

Adverse Perception that Peers have concerning Whistleblowers

Based on two case studies, Van Portfliet (2020) concludes that whistleblowers often experience *stigmatization* by friends and colleagues – where an individual is deemed to possess an attribute that sets them apart from others and is devalued as a person. However, Van Portfliet finds that whistleblowers may respond differently to stigmatization: some accept their fate of being labelled as a whistleblower and treated as

such by their peers for the rest of their working life, whereas others hold the identity "whistleblower" only temporarily and regard themselves as regular employees after some time, and expecting conventional treatment by their peers.

Trevino and Victor (1992) focus on the evaluation and *likeability* of whistleblowers in the aftermath (i.e., acceptability of whistleblowing, ethicality and likeability of the peer reporter). Across their three studies, they find a differing degree of influence concerning the extent to which misconduct threatens the interests of other group members and whether whistleblowing is seen as the responsibility of each individual group member. In both the context of a scenario experiment in an academic setting and in a field survey in a fast-food restaurant, they find that such responsibility positively influences the evaluation of peer reporting as more acceptable and the reporter as less ethical but also more likable. In a third study (a scenario study in a fast-food context), where group interests are at stake, negative emotional reactions to whistleblowing are less negative, and thus the whistleblowing is marginally more acceptable and the whistleblower more likeable in the eyes of peers.

Table 2: Identified Articles about Peer Consequences after Internal Whistleblowing

Author(s), (year)	Context	Country	Sample	Sample size	Method	Peer factor(s)	Category	Main results
Bjorkelo, et al. (2011)	Workplace	Norway	Employees	2,539	Question-naire (quant.)	Exposure to bullying	Adverse actions that peers undertake against whistleblowers	Whistleblowers reported significantly more workplace bullying than non-whistleblowers
Curtis, et al. (2021)	Fraud in organizations	USA	MTurk participant s	Study 1: 256; Study 2: 222	Scenario experiment s (quant.)	Ostracism	Adverse actions that peers undertake against whistleblowers	Intentions to ostracize the whistleblower were significantly higher than intentions to ostracize the wrongdoer
De Maria & Jan (1997)	Fraud in organizations	Australia	Whistle-blowers, superiors	83,214	Survey (quant.)	Ostracism	Adverse actions that peers undertake against whistleblowers	Most whistleblowers experience unofficial reprisals from peers after whistleblowing; most frequent form of unofficial reprisals is ostracism by peers
Jackson, et al. (2010)	Workplace	/	Nurses	18	Qualitative narrative inquiry design (qual.)	Working relationships	Adverse actions that peers undertake against whistleblowers	Whistleblowing had a profound and overwhelmingly negative effect on working relationships; Findings clustered into four themes: (1) Leaving and returning to work, (2) Spoiled collegial relationships, (3) Bullying and excluding, (4) Damaged inter-professional relationships
McGlynn & Richardson (2014)	Academic & financial misconduct, violating rules, rape cover-up	USA	Coaches, faculty members, university staff	13	In-depth interviews (qual.)	Social support	Adverse actions that peers undertake against whistleblowers	Whistleblowers experience reduced social support after whistleblowing; even though peers might express social support in private settings, they avoid doing so in public contexts; whistleblowers' support networks decrease
Raymond, et al. (2017)	Unethical behaviour in organizations	New Zealand	Social workers	10	Interviews (qual.)	Social support, isolation, ostracism	Adverse actions that peers undertake against whistleblowers	Participants report distressing experience of reduced social support, feeling isolated and being ostracized after whistleblowing
Reuben & Stephenson (2013)	/	/	/	68	Laboratory experiment (quant.)	Group in- & exclusion (selection)	Adverse actions that peers undertake against whistleblowers	Participants who previously reported wrongdoing are significantly less likely to be included by the group

Trevino & Victor (1992)	Academic cheating; theft at workplace	USA	Study 1 & 2: students Study 3: employees	Study 1: 478 Study 2: 115 Study 3: 128	Scenarios, field study (quant.)	Likeability	Adverse perception that peers have concerning whistleblowers	Influence of the extent to which misconduct threatens other group members' interests and to which whistleblowing is an individual group member's responsibility on the evaluation of whistleblowers (i.e., acceptability of whistleblowing, ethicality and likeability of peer reporter) differs across three studies; in scenario 2, participants have less negative emotional reactions to whistleblowing, marginally higher acceptability of whistleblowing and likeability of the whistleblower, when group interests are threatened
Van Portfliet (2020)	Sexual harassment; fraud in organizations	/	PhD, senior manager (whistle-blowers)	2	Semi-structured interviews (qual.)	Stigmatization	Adverse perception that peers have concerning whistleblowers	Whistleblowers experience stigmatization by colleagues, but differently respond to stigmatization as the identity "whistleblower" can be temporary and revisable

Adverse Actions that Peers Undertake against Whistleblowers

De Maria and Jan (1997) provide empirical evidence that most whistleblowers experience unofficial reprisals from peers after whistleblowing, most frequently in the form of *ostracism*. Curtis et al. (2021) specifically focus on whether whistleblowers experience *ostracism* as a consequence. They refer to ostracism as the social exclusion or ignorance by peers, and provide empirical support that intentions to ostracize the whistleblower are significantly higher than intentions to ostracize the wrongdoer. These results chime with those of Reuben and Stephenson (2013) where peers select their group members for future cooperation. Those who have previously reported wrongdoing are significantly less likely to be included in the group.

Björkelo et al. (2011) refer to *bullying* as including actions such as harassment, badgering, niggling, freezing out, or offensive teasing, that happen regularly, in which the affected person finds it difficult to defend themselves. In their study, whistleblowers indicate significantly more workplace bullying than non-whistleblowers. Moreover, interviews with whistleblowers conducted by Jackson et al. (2010) revealed the negative effects on *working relationships with other colleagues* as a consequence of whistleblowing. They divide these consequences into four categories: being asked to leave the workplace, damaged collegial relationships (such as barriers created between a whistleblower and their colleagues), bullying and exclusion, and deteriorated inter-professional relationships (e.g., loss of trust).

McGlynn and Richardson (2014) focus on *social support* throughout and in the aftermath of the whistleblowing process. Referring to Goldsmith (2004), they define social support as "what individuals say and do to help one another". They conclude that whistleblowers experience reduced social support in the aftermath. Even though peers might express social support in a private setting, they avoid doing so in a public context. Participants in an interview study in the context of social workers, conducted by Raymond et al. (2017), reported on their distressing experience of receiving reduced social support, and even experiencing isolation.

In conclusion, all the reviewed studies indicate that whistleblowers experience a range of adverse consequences from peers after reporting, whether in the form of being perceived more negatively or through active adverse responses by peers. The results may

show the relevance of addressing more proactively the treatment and value of whistleblowing in the organizational culture to improve the reputation of whistleblowers.

4.5 Discussion

Our systematic literature review complements the existing reviews on whistleblowing by answering the question of how different peer factors relate to whistleblowing. We organize the 27 identified distinct peer factors into a classification of antecedents and consequences (and seven subcategories) of whistleblowing and identify which peer factors have been investigated as moderators and mediators, as summarized in our framework in Figure 1. This framework systematically illustrates that peer factors are diverse and cannot be simplified to one variable, such as the degree of co-worker support. It has to be noted that our framework does not claim to represent a theoretically sound model. Rather, scholars can use our framework in future research to detect categories where empirical results have been mixed and findings worth being replicated in other contexts. Moreover, they may identify and map further peer variables not yet studied, or introduce entirely new categories.

Our findings allow us to derive three main implications and propose future research topics regarding peer influences and consequences (see Table 3, which provides an overview of future research gaps, research paths and exemplary research questions[28]). First, the reviewed studies show that peer factors as moderators, can significantly intensify or weaken the relationship between the main independent variable and whistleblowing. This observation indicates that peer factors interact with further variables. Therefore, investigating whether peer factors can explain previously inconsistent findings on some variables and whether the hypothesized effect only appears under specific social conditions is promising for future research. For instance, in their meta-analysis, Mesmer-Magnus and Viswesvaran (2005) specify that the organizational climate for whistleblowing is less strongly related to actual whistleblowing behaviour than to intentions to blow the whistle. Controlling for peer factors that may significantly influence the relationship with actual reporting might explain this observation. Moreover, combining informal peer factors with the field of formal measures is of particular interest: Accounting for the network of relationships by including loyalty to peers when assessing

[28] While we discuss some of the most salient research issues in the main text, we sketch further research questions in Table 3.

the effectiveness of increasingly formalized whistleblowing programs (and other formal measures) may provide a more comprehensive understanding. In the course of this, it is of interest whether formal specifications on the part of the organization (e.g., codes of conduct) or informal signals about expected behaviour of peers have a stronger influence on whistleblowing. Another aspect that seems worth pursuing is the interaction between individual and peer factors since it is often assumed that individual factors predict whistleblowing less consistently than organizational or situational factors (Mesmer-Magnus & Viswesvaran, 2005; Vadera et al., 2009). For instance, studies on the relationship between whistleblowing and personal morality, often conceptualized by the level of moral development (Kohlberg, 1964; Rest et al., 1999), found mixed support (Vadera et al., 2009). As the concept of moral development is based partly on perceptions of social norms and expectations about one's role, future research could examine whether individuals at certain stages of moral development are particularly influenced by their peers.

Second, the reviewed studies are to some extent not corroborated by a distinct theoretical foundation that explains the results. This indicates that the underlying mechanisms of how and why peer factors affect the whistleblowing decision might require further investigation, i.e., in the form of mediator analyses. On the one hand, we consider this investigation important because peer factors might explain relationships between other variables and whistleblowing. For instance, the anticipation of reprisals from peers could mediate the relationship between a whistleblowing policy (usually prohibiting reprisals) and actual whistleblowing, thereby contributing to a better understanding of how organizations should design policies and which parts to focus on. On the other hand, mediator analyses could advance the understanding of the psychological foundations by which peer factors affect the whistleblowing decision. Theories commonly applied to explain social influences on behaviour such as social identity theory (Tajfel & Turner, 2004), social learning theory (Bandura, 1971), or social comparison theory (Festinger, 1954) might contribute to guiding future research by pointing towards variables that could be considered as mediators. For instance, scholars could ask whether information on peers' behaviour induces concerns about social conformity or social comparison. How do peer factors affect the different step(s) in the whistleblowing process? Does colleagues' behaviour influence the awareness about the wrongdoing and/or the motivation to report? Do social norms in the immediate work

group affect the whistleblowing decision through ethical considerations (e.g., personal norms) or cost-benefit considerations (e.g., by violating the social norm)?

Table 3: List of Proposed Research Gaps, Research Paths & Exemplary Research Questions

Research gaps	**Research paths**	**Exemplary research questions**
Interaction of peer factors with other variables	Interaction of individual and peer factors	• How do peer factors moderate the relationship between personal morality and whistleblowing? • Do peers influence individuals at certain stages of moral development differently? • Do peer factors influence whether individuals with different levels of self-esteem blow the whistle?
	Interaction of situational and peer factors	• How does group cohesiveness towards peers affect the relationship between the severity of the wrongdoing and whistleblowing? • How does emotional closeness to the wrongdoer affect the relationship between the severity of the wrongdoing and whistleblowing?
	Interaction of organizational and peer factors	• How do peer factors moderate the relationship between the organizational climate and whistleblowing? • How does the loyalty to peers moderate the relationship between formal programs and whistleblowing? • Do formal measures (e.g., codes of conducts) or informal signals about expected behaviour by peers have a stronger influence on peers?
	Interaction of different peer factors	• How do antithetical prescriptions of peers affect the whistleblowing decision? • Does trustworthiness of peers influence the relationship of peers‘ ethical behaviour and whistleblowing? • Does emotional closeness to the wrongdoer influence the relationship of advice source and whistleblowing?
Underlying mechanisms of how and why peer factors affect the whistle-blowing decision	Peer factors as explanation for relationship	• How do reprisals from peers mediate the relationship between a whistleblowing policy and whistleblowing? • How does group climate mediate the relationship between organizational climate and whistleblowing?
	Understanding psychological foundations by which peer factors affect the whistle-blowing decision	• Does information on peers’ behaviour induce concerns about social conformity or social comparison with regard to whistleblowing behaviour? • How do peer factors affect the different step(s) in the whistleblowing process? • Does a peer’s behaviour influence the awareness about the wrongdoing and/or the motivation to report? • Do social norms in the immediate work group affect the whistleblowing decision through ethical considerations (e.g., personal norms) or cost-benefit considerations (e.g., by violating the social norm)?

Prevention of negative peer consequences	Promoting whistle-blowing	• How can organizations best communicate the benefits of whistleblowing to all organizational members? • How can organizations link whistleblowing to more positive values? • How can organizations guide peers towards a welcoming whistleblowing attitude?
	Changing the image of whistle-blowing	• How can organizations change the stigma attached to whistleblowing? • How do training and best practices change the image of whistleblowing? • How do employees judge whistleblowers if whistleblowing becomes a duty?

Third, even though peer consequences are relatively less investigated in the reviewed studies than peer influences, we emphasize their importance, because in all reviewed studies, whistleblowers only experience adverse consequences. Therefore, apart from research on how to prevent work-related retaliation, scholars should also focus on preventing negative peer responses. For instance, as proposed by Lewis (2022), the image of whistleblowers should be changed from being seen as acting against the norms of the group to being decent employees following a duty. However, this may entail that employees will be held liable for ethics at work, which eliminates individual responsibility and moral autonomy (Tsahuridu & Vandekerckhove, 2008). Future research may therefore examine whether and how policies that regard non-reporting of wrongdoing as a violation can increase internal whistleblowing, or whether this kind of policy backfires because employees are perceived to be treated unfairly as long as whistleblowing is still regarded as an act that requires moral courage above average standards. Scholars could also investigate how organizations can communicate the benefits of whistleblowing to all organizational members and link whistleblowing to more positive values, such as organizational citizenship behaviour (Organ, 1988) or in- and extra-role behaviour[29] (Bjørkelo & Macko, 2012). Stressing the benefits and improvements of an organization's service, products or processes achieved through whistleblowing might provide a first step in transforming the image of whistleblowing (Bjørkelo & Macko, 2012). However, stigma cannot only be transformed by merely changing the image, because reactions can also be triggered by underlying beliefs, which

[29] While in-role behaviour is a requirement and part of the duties an employee has to fulfil, extra-role behaviour describes discretional behaviour beyond the daily expectations of employees (Bjørkelo & Macko, 2012).

need to be transformed accordingly. Changing these beliefs is not easy or straightforward, however, due to deeply rooted social norms (e.g., not snitching on a colleague).

From a methodological point of view, the reviewed studies have examined whistleblowing intentions mainly through surveys involving self-reporting and hypothetical contexts. This may be problematic as predictors of intentions are often distinct from predictors of actual whistleblowing (Mesmer-Magnus & Viswesvaran, 2005). Likewise, interviews with actual whistleblowers may be restricted in their validity due to potential selection effects and the lack of control groups. Hence, complementing empirical approaches with observational field data and behavioural experiments would be helpful. Although scholars should place emphasis on observing and describing first-hand actual behaviour in organizations, behavioural experiments in the laboratory or in the field are still noteworthy for several reasons. Systematically studying peer-related whistleblowing with observational field data may be challenging because both witnessing and reporting wrongdoing tend to be rare and are often confounded by other factors. For example, it would be extremely difficult to examine, in a field setting, the causal relationship between policies and whistleblowing behaviour, because this causality can potentially be affected by the fear of ostracism.[30] By contrast, by exercising tight control over confounding variables and the decision environment (Hauser et al., 2017), laboratory experiments provide a more suitable empirical test environment for examining and clearly identifying causal relationships derived from theories. This might be of particular relevance when theories and social relationships are not tied to special organizational structures, circumstances and commodities that can be simplified and simulated in a more abstract decision environment. Concerning the examination of the potential research question stated above, laboratory experiments thus allow creating controlled conditions with and without the fear of ostracism, hence investigating the moderating effects of ostracism on the relationship between policies and whistleblowing. Likewise, in respect of the important study of mediator analyses, researchers are usually interested in the interaction of a few, specific variables. Thus, to detect the underlying mechanisms of a potential outcome and the interdependence of the independent variables, it is more important to control for possible social interactions, ties and group dynamics, as well as

[30] Analysing this causal relationship with field data could be difficult because the decision of the organization in favour of a policy is not the result of an exogenous and randomized process, but a deliberate decision (i.e., self-selection). Further, information about the prevalence, severity and form of ostracism is seldom reported and may not even be known.

their controlled, exogenous variation (in the laboratory more than in the field). Research on interaction effects can also be conducted using observational field data. However, longitudinal data on group variables such as dynamics may not be available or cross-sectional data may not be able to duplicate the controls that are necessary for mechanism testing.

The review is based on two reviewers who independently decided upon the inclusion and exclusion criteria and the cluster of categories following the given review methodology. Results were synthesized after each step with the purpose to structure the results. A potential limitation is that we only included studies published in peer-reviewed journals, and do not assess the quality of the reviewed studies in terms of, for example, their differing empirical value, the nature and size of the sample, or the strength of the findings. An additional difficulty arises from the different publication contexts (e.g., business, nursing, sports), which renders uniform weighting of the quality of studies more difficult.

Nevertheless, our review seeks to raise awareness about the conflicting loyalties that potential whistleblowers experience in the organization and which impact their willingness to report wrongdoing. Further research on how to resolve such conflicting interests by addressing peer factors alongside and combined with individual, situational and organizational determinants would seem to be beneficial in helping organizations achieve their goal of guarding themselves against wrongdoing.

CHAPTER 5 | "I don't believe that you believe what I believe": An Experiment on Misperceptions of Social Norms and Whistleblowing

Behnud Mir Djawadi Sabrina Plaß Sabrina Loer

Social norms fundamentally shape economic decision-making, yet individuals often systematically misperceive what others think and do, potentially leading to suboptimal social outcomes. We examine how such misperceptions affect behaviour and whether correcting them can induce a behavioural change, using whistleblowing as an application. Through survey data and an incentivized laboratory experiment, we demonstrate that while a majority of individuals (>75%) privately support whistleblowing, almost half (45.92%) misperceive the majority's view. Both personal normative beliefs and normative expectations strongly predict whistleblowing behaviour. Their alignment is particularly noteworthy: individuals who personally support whistleblowing are more likely to report misconduct when they believe others share their views. A social information intervention revealing the true distribution of peer support affects subgroups differently: while it increases whistleblowing behaviour among individuals who already personally favour reporting misconduct, there is no effect among those who are personally resistant to it. Still, given the relatively low cost of such social information interventions, they offer an economically viable means of achieving behavioural change in at least some of the targeted individuals.

Keywords: *Social Norms, Misperceptions, Normative Expectations, Personal Normative Belief, Whistleblowing*

JEL Codes: C92, D01, D83, D9

5.1 Introduction

Social norms play a fundamental role in shaping human behaviour and economic decision-making (Bicchieri, Dimant, Gelfand, et al., 2023). They act as informal rules that guide social interactions across various domains, motivate behaviours, such as pro-social behaviour (Gächter et al., 2017; Krupka & Weber, 2009) and charitable giving (Agerström et al., 2016), and even help reduce discriminatory behaviour against out-group members (Barr et al., 2018). While economic decision-making has long been assumed to be primarily driven by monetary incentives (e.g., Gneezy et al., 2011), growing evidence demonstrates that people’s actions are also substantially influenced by what they believe others think and do (Bicchieri, 2006).

However, emerging evidence suggests that individuals often systematically misperceive the prevalent social norms in a given situation. For instance, Bursztyn et al. (2020) document that while the vast majority of Saudi men privately support women working outside the home, they substantially underestimate this support among their peers. This misperception contributes to persistently low female labour force participation despite potential economic gains from greater inclusion. Thus, as the outcome of misperceived social norms is neither the result of individual preferences nor of social consensus, the societal and economic costs can be substantial and entail undesired social outcomes. The misperception of social norms can prevent change to existing practices, even if such change would bring about mutually beneficial transactions and welfare-enhancements (Bursztyn & Yang, 2022).

Our research aims to contribute to the understanding of the behavioural consequences of these misperceptions, and investigates if the correction of misperception can induce behavioural change. In a comprehensive meta-analysis of the recent empirical literature, Bursztyn and Yang (2022) find that misperceptions are widespread across domains, highly asymmetric, and are substantially larger when concerning out-group members compared to in-group members. Moreover, they establish that individuals’ own attitudes and beliefs have a strong positive association with their perceptions of others’ attitudes, suggesting deep interconnections between personal views and social expectations. However, we currently lack a more nuanced understanding into how personal views and (misperceived) social expectations translate into actual behaviour, and, even more interestingly, if attempts to correct such beliefs can induce behavioural change and for

whom. Although existing research demonstrates that the correction of false normative expectations can increase the desired behaviour (Andre et al., 2024; Bursztyn, González, et al., 2020; Wenzel, 2005), it remains largely unanswered whether this also applies to situations where the desired behaviour carries a personal cost for the decision-maker, and where the alternative behavioural option is relatively risk free. In these settings, interventions that seek to re-calibrate beliefs may still be insufficient to change behaviours in the direction of the corrected beliefs. Moreover, although prior studies also investigated the heterogenous effects of norm interventions for different subgroups (e.g., Andre et al., 2024; Bursztyn et al., 2020), they did not explicitly focus on how the correction of normative expectation is related to personal normative beliefs.

We therefore conduct a comprehensive empirical analysis to examine a) the actual behavioural consequences of both personal views and (misperceived) social expectations, and b) whether/ how a social information intervention that corrects the expectations also changes behaviour.

We decided to use a whistleblowing[31] framework, for the following reasons. First, the whistleblowing decision environment exhibits structural features that theoretically facilitate the emergence of norm misperceptions: (a) private attitudes about reporting misconduct are rarely openly discussed in organizations, creating systematic information gaps (Knoll & Van Dick, 2013), (b) the potential whistleblower faces significant perceived social costs (De Maria & Jan, 1997; Reuben & Stephenson, 2013), and (c) the low observability of actual whistleblowing incidents makes it difficult for individuals to accurately capture the true distribution of attitudes in their reference group (Dyck et al., 2010). Second, the whistleblowing context exhibits a particularly interesting paradox: while survey evidence consistently shows majority support for whistleblowing in principle (e.g., pointing to shared ethical standards and professional responsibilities) (Keenan, 2007; Skivenes & Trygstad, 2016), actual reporting rates remain notably low (Butler et al., 2020). The discrepancy between stated preferences and observed behaviour hints to the presence of systematic misperceptions about peers, which may explain the continued persistence of non-reporting. Third, whistleblowing is costly (Miethe &

[31] We define whistleblowing as “the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action” (Near & Miceli, 1985: 4). The focus of the study lies on internal whistleblowing, the timely detection and correction of misconduct within the organization that avoid costs arising from reputational damage caused by external disclosures (Lee & Fargher, 2013)

Rothschild, 1994; Rothschild, 2008). This presents a particular challenge for the effectiveness of interventions seeking to correct beliefs, because achieving behavioural change may require a more persistent and comprehensive strategy than merely providing information about what others truly believe.

Hence, we use a whistleblowing application to examine the misperception of social norms, specifically asking how potential whistleblowers personally think, what they expect others to think about whistleblowing, and how this relates to their own behaviour. If there are misperceptions in the whistleblowing context, does correcting the beliefs about others increase whistleblowing behaviour?

Our empirical approach employs a survey and an incentivized laboratory experiment. For both methods, we use the conceptualization of Bicchieri (2006) to operationalize personal attitudes (=personal normative beliefs) and the perception about others (=normative expectations), as it is well suited to economic experiments (Dubreuil & Grégoire, 2013) and widely used in the economic literature.[32] Personal normative beliefs are an individual’s own beliefs regarding the right course of action in a given situation while normative expectations are expectations about what the majority of people in a given reference group think is the right thing to do (Bicchieri & Xiao, 2009). In contrast, empirical expectations are expectations about how most people in the reference group behave in a certain situation (Bicchieri, 2006). For our research objective, we measure personal normative beliefs and normative expectations with the latter described as second-order beliefs, meaning that normative expectations represent beliefs about others’ personal normative beliefs. We ran our survey on the platform Prolific, asking employees who have witnessed wrongdoing in their workplace about their personal normative beliefs and their perception about their colleagues’ beliefs about whistleblowing. While the design of our survey study does not allow us to detect actual misperceptions – because we cannot verify the majority view in the organization of the surveyed employees – we were able to identify a discrepancy between participants’ personal normative beliefs and their normative expectations, revealing the potential existence of a misperception. Our incentivized laboratory experiment conducted with a student sample allowed us to systematically elicit both personal normative beliefs and normative expectations, and

[32] Other conceptualizations of norms do exist, for example, Cialdini et al. (1990) and Rimal & Real (2003). What they have in common is the inclusion of a descriptive (what others do) and a prescriptive (what one should do) component (Dannals & Miller, 2017).

correlate them with real whistleblowing behaviour. Our experimental results show that most participants (77.55%) in a baseline treatment personally support whistleblowing. However, almost half of the subjects (45.92%) misperceive the predominant personal normative belief of their peers, and thus show a misperception. In terms of actual behaviour, we observe that holding personal normative beliefs in support of whistleblowing is positively associated with whistleblowing behaviour, however, we observe less whistleblowing for participants whose normative expectations run against their personal belief. To correct for these misperceptions, we used in our second treatment a social information intervention[33] that communicated to participants the true majority’s belief prior to the decision. Compared to the baseline treatment we do not see an increase in whistleblowing behaviour across the whole sample. However, significantly more subjects with a personal normative belief in favour of whistleblowing report the observed wrongdoing if they get the social information intervention.

Our study contributes to different strands of the literature. First, similarly to social norm misperceptions in other contexts, such as of women’s participation in the labour market (Bursztyn, González, et al., 2020), tax evasion behaviour (Wenzel, 2005) or actions against climate change (Andre et al., 2024), we find that wrong beliefs about others have the effect of discouraging desired behaviour, in our case, whistleblowing. We find that if the majority view is believed to be against whistleblowing, whistleblowing is less likely to be observed, even if an individual personally supports it. Consistent with social norm frameworks (e.g., Krupka & Weber, 2013), we find that personal normative beliefs and normative expectations are strong predictors of whistleblowing behaviour. Further, both concepts complement each other in the prediction of behaviour: whistleblowing is more likely observed if both personal normative beliefs and normative expectations favour whistleblowing than if they stand in conflict.

Second, while there has been some positive evidence of social information interventions, for example, to reduce littering (Cialdini et al., 1990) or using free plastic bags in supermarkets (De Groot et al., 2013), normative information did not work as well in various other important domains, such as increasing return rates in a trust game (Bicchieri et al., 2021), enhancing tax compliance (Hallsworth et al., 2017), or encouraging honesty (Dimant et al., 2020). Our result of not finding a significant increase

[33] Also termed (among others) norm-based interventions (Miller & Prentice, 2016), social norm messages (Bhanot, 2021), norm cues (Cialdini et al., 1990) or social norm nudges (Bicchieri & Dimant, 2022a).

in whistleblowing behaviour across all experimental subjects, would suggest support for this part of the literature. However, as in this second treatment we also elicited personal normative beliefs, we can conclude that our intervention affects subgroups differently. In contrast with Andre et al. (2024), we find that individuals were not responsive to the intervention if they were personally opposed to the behavioural option. Hence, persuading those individuals in the whistleblowing context needs more than bridging an information gap. However, we observed a weakly significant 12% increase in the number of individuals if they personally supported the behavioural option promoted by the intervention. Our finding therefore aligns with the results of the literature which states that interventions revealing the truth about others’ personal normative beliefs helps especially those individuals who are already personally in favour of the targeted behaviour (e.g., Bursztyn et al., 2020; Wenzel, 2005).

Third, we identify whistleblowing as another economic domain where misperceptions about others exist. Attitudes among employees are actually more positive about whistleblowing than employees themselves assume. As fearing retaliation from colleagues in the aftermath of reporting is among the most frequently cited reasons that deters people from whistleblowing (e.g., Rothschild & Miethe, 1999), correcting the misperception they hold about others (i.e., that the majority is not in support of whistleblowing) has the potential of alleviating the concerns about potential social sanctions that could follow the act of blowing the whistle, at least for those already favourably inclined towards it.

The rest of the paper proceeds as follows: in Section 2, we derive our predictions. In Section 3, we briefly describe our survey conducted with real employees and highlight the main results. In Section 4 we describe our experimental design, present the results in Section 5, and a conclusion in Section 6.

5.2 Predictions

Misperceptions in the Whistleblowing Context

Several empirical studies provide first indications that a majority of people personally support whistleblowing. In a field experiment at a large financial services firm, Burks and Krupka (2012) show that employees and corporate leaders overwhelmingly consider whistleblowing appropriate in clear misconduct cases. Dyck et al. (2010) reveal that a

reluctance to report misconduct stems primarily from concerns about retaliation and career implications rather than personal opposition. This finding is also observed across various industries (e.g., Near et al., 2004; Rothschild & Miethe, 1999). We further posit that, while most individuals privately favour whistleblowing, they are less certain whether this view is equally shared by others, notably among work colleagues, thereby miscalibrating the true beliefs of others. Several mechanisms may drive this misperception. First, deriving normative expectations from the frequency by which whistleblowing occurs may misrepresent the true distribution of personal normative beliefs, as many potential whistleblowers may suppress their intentions due to fear of retaliation. Second, while the negative consequences of whistleblowing are often more visible to peers, private support remains largely unobserved, potentially resulting in the phenomenon of pluralistic ignorance, where individuals incorrectly believe they hold a minority view (Bursztyn, González, et al., 2020; Munsch et al., 2014). Finally, the rarity of whistleblowing, coupled with frequent peer sanctions (Mechtenberg et al., 2020) reinforces – openly or more subtly – perceptions of widespread disapproval. These misperceptions may manifest as, on the one hand, an underestimation by the supporters of whistleblowing, and/or, on the other, an overestimation by its opponents. We maintain a conservative position, considering the potentially substantial misperceptions in both groups without specifying relative differences in magnitude:

> *Prediction 1: There is a general misperception about the majority’s personal normative beliefs in form of an underestimate of the support for and/or an overestimate of the opposition against whistleblowing.*

Personal Normative Beliefs, Normative Expectations, and Whistleblowing Behaviour

Drawing on conventional utility frameworks from the (social) norm literature (e.g., Krupka & Weber, 2009) and related empirical evidence that we mention below, we predict how personal normative beliefs and normative expectations independently and jointly are factors that influence whistleblowing behaviour. Burks and Krupka (2012), and Bašić and Verrina (2023), emphasize that personal normative beliefs operate separately from social norms, with both independently affecting the utility evaluation of an individual’s actions. We predict that when an individual’s personal normative belief is pro-whistleblowing, this should increase their consideration of the utility from and the

likelihood of whistleblowing behaviour, ceteris paribus. Similarly, assuming that individuals care about the opinions of others, as is strongly suggested by the whistleblowing literature (e.g., Mir Djawadi et al., 2023), the utility from whistleblowing actions should be higher when individuals believe others to endorse rather than oppose whistleblowing, making such behaviour more likely. Related empirical literature supports these conjectures: individuals with high personal ethical standards are more likely to express their intention to blow the whistle (e.g., Liyanarachchi & Newdick, 2009), while factors that discourage whistleblowing include the fear of reprisal from peers (Iwai et al., 2021; Rennie & Crosby, 2002), the threat of being excluded from a group after whistleblowing (Reuben & Stephenson, 2013), and loyalty towards peers (e.g., Goddiksen et al., 2021; Taylor & Curtis, 2010). We summarize our two predictions as follows:

> *Prediction 2: A personal normative belief in support of whistleblowing is positively related to whistleblowing behaviour compared to a personal normative belief in support of remaining silent.*
>
> *Prediction 3: Normative expectations in support of whistleblowing are positively related to whistleblowing behaviour compared to normative expectations in support of remaining silent.*

Another implication derived from these frameworks is that the utility of an action is considered higher if personal normative beliefs and normative expectations are aligned with each other than if they stand in conflict with each other. We therefore hypothesize that individuals who personally consider whistleblowing to be the correct action are more likely to blow the whistle when their normative expectations are also in support of whistleblowing than when they are against it. Burks and Krupka (2012) find that misalignment between personal and group norms correlates with lower job satisfaction, and reduces the probability of telling the truth. Other related research shows that ethical conflict between personal and organizational values predicts turnover (Schwepker, 1999), while alignment of the ethical views of the individual and their group increases organizational commitment (Ambrose et al., 2007).

> *Prediction 4*: *For those with a personal normative belief in support of whistleblowing, believing that peers also support whistleblowing (normative expectations aligned with personal beliefs) is positively related to whistleblowing*

behaviour compared to believing that they are against it (misaligned expectations).

Social Information Intervention to Correct Normative Expectations

If our previous predictions hold, we should find that both personal normative beliefs and normative expectations are relevant factors in an individual’s decision-making process, and that the true normative climate is overall in support of whistleblowing (i.e., the majority of individuals’ personal normative belief is pro-whistleblowing). According to Bicchieri (2006), communicating the majority’s true personal normative belief would make normative expectations vis-à-vis whistleblowing more salient. This has two implications: first, for the individuals who hold personal normative beliefs in support of whistleblowing, but who think others are against it, it would resolve the discrepancy between their personal normative belief and their normative expectation of others, and thus lead to an increase in whistleblowing behaviour. Second, for the individuals who hold a personal normative belief against whistleblowing, and think others are also against it, information about the majority’s actual attitudes would create a discrepancy between their personal normative belief and their normative expectation. Consequently, fewer individuals may decide to remain silent compared to a situation where information about the true share of personal normative beliefs is unknown. Our last prediction, therefore, is:

Prediction 5: Information, prior to the decision of whistleblowing, that the true personal normative belief held by the majority is in support of whistleblowing increases whistleblowing behaviour.

5.3 Survey

Concurrently to conducting our experiment, we ran a survey via the platform Prolific[34] in May 2023 to examine whistleblowing perceptions among employees, and to identify potential indications of misperceptions through discrepancies between personal normative beliefs and normative expectations in a real-world setting. In total, we recruited 100 participants who completed a 7-minute-long survey for a fixed payment of £1.40. Participants came from different countries of residence (e.g., Spain, UK, Poland,

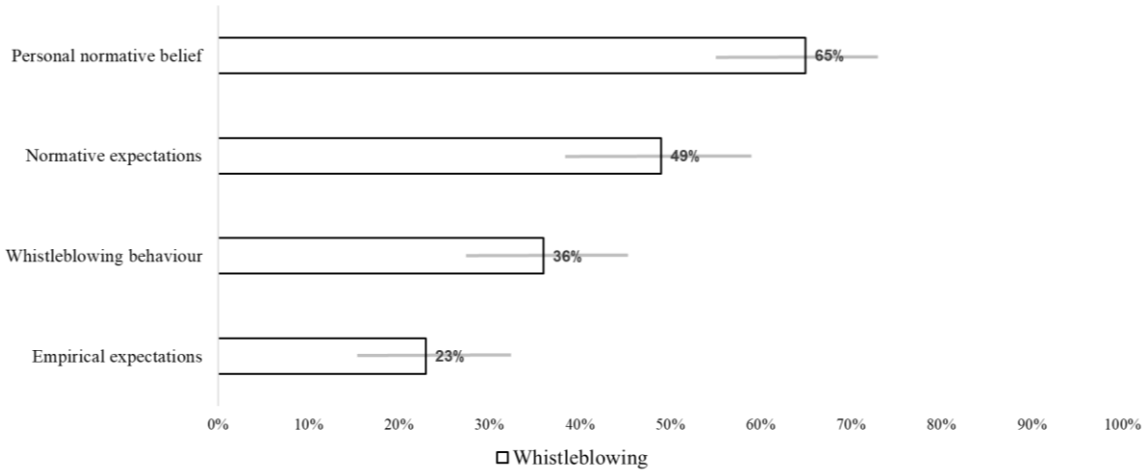
[34] The main rationale of choosing Prolific was the possibility to survey real employees who witnessed wrongdoing in their organization. For a more detailed discussion of the advantages of different crowdworking platforms, see, e.g., Palan and Schitter (2018) or Peer et al. (2017).

Germany), organizational positions (e.g., upper management, consultant, administrative staff) and sectors (e.g., business management, education and training, retail, information technology). The demographic composition was 50% female, 49% male, and 1% binary, with a mean age of 30 years. Besides being employed at an organisation at the time of the study, an essential participation requirement (implemented via Prolific’s screening protocol) was for participants to have observed wrongdoing in the workplace.

Participants received the following instructions: “Please think of the organisation where you observed the wrongdoing. Please also think of your colleagues in that organisation.” Participants should then indicate how they behaved in the situation and specify some characteristics of the observed wrongdoing. We captured participants’ personal normative beliefs by asking them to, “Indicate which behaviour (independently of what you actually did) was for you personally the right thing to do at that time: a) not report the colleague, b) report the colleague”. Then, we asked them to indicate what they think their colleagues at that time would have stated as the right thing to do (normative expectations). For the sake of completeness, we also measured empirical expectations by asking what participants think their colleagues would have done in the situation. Additionally, we asked them about their motivation, organisation- and job-related variables, and demographics. Two attention checks were also included.

The survey’s main objective was to investigate potential discrepancies between personal normative beliefs and normative expectations regarding whistleblowing, and to conduct preliminary analyses of their respective relationship to actual behaviour. We therefore remain brief in our analysis here and refer for a more detailed version that can be found in the Appendix A of Chapter 5 (p. 221). As illustrated in Figure 1, the majority of subjects (65%) held the personal normative belief that reporting was the right course of action. However, only 49% of the subjects held the normative expectation in support of whistleblowing, that is, slightly less than half of the subjects thought their colleagues believed whistleblowing to be the right course of action. Even fewer subjects (only 23%) expected their colleagues to blow the whistle in that situation. The systematically lower proportions for both normative and empirical expectations, compared to personal normative beliefs, suggest a potential discrepancy between individuals’ personal view on whistleblowing and their perceptions of their colleagues’ views.

Figure 1: Share of Survey Participants with Personal Normative Belief, Normative Expectations, Behaviour, and Empirical Expectations in Support of Whistleblowing (n = 100)



Note: Grey lines indicate 95% confidence intervals.

Moreover, we observe a discrepancy between personal normative beliefs and normative expectations regarding whistleblowing in 34% of participants. For the majority of the latter, the discrepancy goes in the presumed direction, that is, 25% of all participants held the personal normative belief that whistleblowing was the right thing to do when observing wrongdoing in their organisation but expected that the majority of their colleagues believe that not reporting the wrongdoing was the right thing to do. The results also provide an initial insight into the relationship between personal normative beliefs, normative expectations, and whistleblowing behaviour. Both personal normative beliefs and normative expectations seem to have a significant positive relationship with self-reported whistleblowing behaviour ($\chi^2 = 17.582$; $p < 0.001$ and $\chi^2 = 9.408$; $p = 0.001$)[35]. The relationship between normative expectations and whistleblowing behaviour remains consistent if considering only participants who believe whistleblowing to be the right thing to do ($\chi^2 = 5.726$; $p = 0.009$).[36]

5.4 Experiment

Basic Experimental Set-up

To systematically analyse the influence of normative beliefs and expectations on whistleblowing perceptions and behaviours, we conduct a two-part incentivised

[35] We used one-sided tests throughout the study as our hypotheses are directional.
[36] The relationship between normative expectations and whistleblowing is statistically insignificant if we only consider the 35 participants that hold a personal normative belief against whistleblowing ($\chi^2 = 1.136$; $p = 0.144$).

laboratory experiment. Table 1 presents the experimental timeline and stage descriptions. For the experimental design we took inspiration from two sources: the whistleblowing design by Bartuli et al. (2016)[37], and the two-step process for eliciting personal normative beliefs and normative expectations established by Bicchieri and Xiao (2009). In the first part of the experiment, we randomly matched two subjects to form a team that plays together for three consecutive rounds. In each round, the team collaborates on a joint real-effort task to earn an income. Each of the two players is given 2 minutes to count how many times the digit “7” occurs in a matrix of 300 numbers. The counts of the two players are added together, and if the difference between this sum and the actual count of the digit “7” is not greater than four, we deem the task to be successfully solved, and both players receive a payment of 5 tokens. If the task is not successfully solved, neither player receives a payment for the respective round. Participants are informed after each round whether their team solved the task correctly and receive the payment. Following these three rounds, we introduce a situation involving wrongdoing and whistleblowing. We assign one member of each team the role of Player A and the other the role of Player B. Player A is given a donation budget of 5 tokens and the task of forwarding it to a named charity[38]. However, Player A has the option to keep part of the money (2 tokens) to themselves, in which case we refer to it as the embezzled part of the donation.[39] The team member in the role of Player B can react to the embezzlement of Player A either by *remaining silent* about the embezzlement or by *blowing the whistle* on it[40]. The strategy method is used to capture each Player B’s reaction to the potential embezzlement of Player A. This involves Player B being asked how they would act if Player A were to keep part of the donation budget, at the same time as Player A is making their decision, but without Player B knowing yet what Player A decides. The consequences of these decisions are structured as follows: if Player A forwards the full donation, Player B’s decision remains without consequences. However, if Player A does embezzle part of the

[37] Unlike them, however, we refrain from framing the experiment in an organizational context and opt for a neutral framing.
[38] The charity is called *GoAhead!* (http://goahead-organisation.de). On a scale from 1 to 7, participants in our experiment rate the worthiness of support of the organisation with an average score of 6.08. Therefore, we assume that participants overall consider the organization to be worth supporting.
[39] We included a corresponding control question to ensure that the situation was perceived as misconduct. On a scale from 1 to 7, on the question of whether the embezzlement of the donation represents a wrongdoing, participants in the role of Player B indicate on average a score of 5.27 (5.61 in Treatment 1; 5.03 for Treatment 2). This suggests that participants in the role of Player B tend to view the embezzlement of the donation as a wrongdoing.
[40] We deliberately adopted neutral language in the instructions the participants received, for example by using the terms “ignoring”, “reporting”, “keeping part of the donation”, etc.

money, Player B’s decision is implemented: if they said that they would remain silent about the embezzlement, Player A’s payoff is increased by 2 tokens, and the charity loses these 2 tokens, while Player B’s earnings remain unaffected. If Player B reports the embezzlement, however, the charity receives the full amount of 5 tokens, while both players’ payoffs are reduced as follows: 2 tokens are deduced from Player A’s payoff (as punishment), and 1 token from Player B’s payoff, to account for the cost of whistleblowing (e.g., retaliation or ostracism). When making their decision, Player A knows about Player B’s option to blow the whistle, and both are aware of the consequences of all possible combinations of actions.

After both players have made their decision, we elicit the personal normative beliefs, the normative expectations, and the empirical expectations from both players (embezzlement for Player A, whistleblowing for Player B). This elicitation differs somewhat between our two treatments and is therefore described in more detail in the “Treatments” section. Participants know about the elicitation from the instructions, but do not know yet what questions will be asked.

After the elicitation of beliefs and expectations, participants are informed about the decisions regarding the donation budget and their respective consequences. This stage concludes the first part of the experiment.

The second part of the experiment is designed to account for possible social sanctions in the aftermath of the wrongdoing/whistleblowing decision. Specifically, team members are randomly matched again, but with someone who previously had the same player type role (A or B), that is, either being responsible for the donation budget, or for the (potential) whistleblowing. In the new team of two, we assign two new roles: one subject is assigned the role of Player 1, the other of Player 2. Player 1 must decide whether or not to work with the new team member, based on Player 2’s previous decision (donation/embezzlement or silence/whistleblowing). Once again, we apply the strategy method.[41] If Player 1 decides to work in a team with Player 2, both players solve the same real-effort task together, just as in the first part of the experiment, and they can increase their payoffs by 5 tokens. However, if Player 1 excludes Player 2 from the team, Player 1

[41] For a pair of previous Players A, Player 1 is asked how they would decide if they were matched with a Player 2 who previously donated the money, or with a Player 2 who previously kept part of the money; for a pair of previous Players B, Player 1 is asked how they would decide if matched with a Player 2 who previously indicated they would ignore the wrongdoing, or with a Player 2 who previously indicated they would report it.

has to solve the task individually, with a payment of only 4 tokens if the task is solved correctly. As Player 2 is excluded from the task, they cannot earn any money in this round. Right from the beginning of the experiment participants are made aware about the potential choices, and the associated consequences, of this second part.

Table 1: Description of all Stages of the Experiment

Timeline of a session.	
Stages	**Description**
1	Participants welcomed; **session starts.**
2	**General Instructions** read silently. Instructions provide all the information about Part I and Part II of the experiment.
3	**Part I** of the experiment starts.
4	**Real Effort** Team Task (3 rounds).
5	**Feedback** about success of **Real Effort Task** (after each round).
6	Participants become either **Player A or B.**
7	**Donation** decision (Player A); **Reporting** decision (Player B, strategy method).
8	**Elicitation** of personal normative beliefs and normative expectations (T1, incentivized) or empirical expectations (T2, incentivized).
9	**Feedback** about **Donation** and **Reporting** decision and the associated consequences**.**
10	**Part II** of the experiment starts; participants become either **Player 1 or 2**.
11	Player 1 decides whether to do the **Team Task** or the **Individual Task** (exclude Player 2) based on the decision Player 2 took in Part I (strategy method).
12	**Decision** of Player 1 **implemented** based on behaviour of Player 2 in Part I.
13	Participants are **informed** about their **total payoffs**, the **donation** generated for the charity and whether they successfully assessed normative expectations/empirical expectations.
14	**Experiment ends; Questionnaire starts.**
15	**Session ends**.
16	Participants are **paid** anonymously accordingly to their earnings in the experiment.

Treatments

We form two treatments that differed in the way the normative measures and presented information about the majority’s personal normative beliefs were elicited. In *Treatment 1*, we follow the two-step approach by Bicchieri and Xiao (2009) for the elicitation of personal normative beliefs and normative expectations[42]. In the first step, we ask Player B about their personal normative belief regarding which of the two behavioural options – remaining silent or blowing the whistle – they personally consider the morally right thing to do. In the second step, subjects can earn an additional token if they correctly assess the majority opinion of other subjects in the role of Player B in the same experimental session regarding the previous question (normative expectations on whistleblowing). This enables us to detect individual misperceptions, i.e., if a participant incorrectly assesses the majority opinion regarding whistleblowing. Finally, we ask subjects in *Treatment 1* about their beliefs (empirical expectations) about what they think the majority of Players B (actually) decided in the situation (remain silent vs. blow the whistle).[43]

Treatment 2 modifies this structure by providing Player B with explicit information about the majority’s normative belief on whistleblowing prior to their decision. This information is derived from data collected in *Treatment 1,* and participants are informed, through a sentence on the screen, that “In previous sessions of the same experiment with students from the same University, the majority of participants indicated that the morally correct behaviour for them personally would be to report the behaviour of Player A.” Thus, participants in *Treatment 2* do not need to form their own normative expectations but can rely on the information provided to them. Accordingly, in *Treatment 2,* we only ask Players B which of the two behavioural options – remaining silent or blowing the whistle – they personally consider the morally right thing to do (personal normative belief). Additionally, we ask them what they believe the majority of Players B

[42] We preferred this method to the other widely-used norm elicitation method established by Krupka and Weber (2013) for the following reasons: to analyse discrepancies in beliefs, and to establish the informational basis of the social information intervention, we needed to measure both first-order beliefs (personal normative beliefs) and second-order beliefs (normative expectations) which the opinion matching method addresses better than the Krupka/Weber method. Additionally, the elicitation of personal normative beliefs enables us to analyse different subgroups, and how they are influenced by the social information intervention. Lastly, the literature (e.g., Görges & Nosenzo, 2020) suggests using the opinion matching method if one can safely assume the absence of a widely accepted social norm, as is the case in the whistleblowing context.

[43] The same questions were asked of Player A regarding the forwarding or withholding of the donation and the behaviour of other Players A.

in the session have (actually) decided in the situation (empirical expectations). To maintain equal potential payoffs in both treatments, participants in *Treatment 2* can earn an additional token for the second question if they correctly assess the majority’s behaviour.[44] Participants only find out at the end of the experiment whether they correctly assessed the normative expectation (Treatment 1) or the empirical expectation (Treatment 2).

Procedure

We conducted the experiment between May 2023 and November 2024 in an economic research laboratory of a German university. A total of 396 undergraduate students from various fields of study were randomly selected from the university’s experimental subject pool managed by the online recruiting system ORSEE (Greiner, 2015). The experiment was programmed using oTree (Chen et al., 2016). We conducted a total of 23 sessions, of which 16 took place in the lab and 7 online, with each subject participating in only one session. In the 16 lab-based sessions, the registered participants were randomly seated in cubicles in front of a computer screen, where they could not be observed by others. In the 7 online sessions, registered subjects received a link to the experiment via Zoom, to which they were logged in with an alias name so that they could not be identified by other participants. Before the experiment started, all participants received written instructions explaining the procedure and rules of the entire experiment (for the instructions, Appendix B of Chapter 5, p. 225). After completing the experiment, participants answered an additional questionnaire capturing socio-economic demographics (including age, gender, and field of study), questions about donation behaviour, social comparison orientation and manipulation checks. The experimental sessions lasted 45 minutes and participants earned on average 8.93 € (incl. a show-up fee of 2.50 €; with 1 token of the experimental currency worth 0.50 €).

[44] Again, we ask the same questions of Player A regarding the forwarding or withholding of the donation and the behaviour of other Players A. Moreover, Players A in Treatment 2 also receive the information about the majority’s personal normative beliefs with regard to the donation/embezzlement decision based on the data of Treatment 1.

5.5 Results

Given the focus of our study, in the following analysis we exclusively report the results for participants in the role of Player B.[45] We first provide information about the composition of the sample, the performance in the real-effort task, and the participants excluded from the team task in part 2 of the experiment for each treatment. Then, in the hypothesis testing, we follow the order of our predictions, focusing on Treatment 1 for predictions 1 to 4, and presenting the results of Treatment 2 for addressing prediction 5. Lastly, we corroborate our results through a regression framework.

Descriptive Data

In total, 198 participants were assigned the role of Player B[46], with 98 in Treatment 1 and 100 in Treatment 2. Table 2 presents the descriptive statistics of the sample per treatment among Player B participants. There were no significant differences across treatments in terms of age (Mann-Whitney U Test: $z = -0.090$, $p = 0.928$), gender (Chi-Square Test: $\chi^2 = 1.108$, $p = 0.575$), field of study (Chi-Square Test: $\chi^2 = 1.303$, $p = 0.728$), performance in the real-effort tasks in Part 1 (Mann-Whitney U Test: $z = -1.441$, $p = 0.150$) or Part 2 (Mann-Whitney U Test: $z = -1.109$, $p = 0.267$), and the Social Comparison Orientation Scale by Gibbons and Buunk (1999; as a measure for people’s tendency to compare themselves to others) (Mann-Whitney U Test: $z = 1.417$, $p = 0.156$).

Moreover, there is no difference between the two treatments in the number of participants in the role of Player B who are excluded from the task in the second part of the experiment (Chi-Square Test: $\chi^2 = 0.2886$, $p = 0.591$). More precisely, in Treatment 1, 12 Players B (24.49%) are excluded from the team task, of which six are previous whistleblowers and six are Players B who remained silent about the observed embezzlement. In Treatment 2, of the 10 Players B (20.00%) that are excluded from the team task, seven remained silent about the embezzlement, and three reported the embezzlement.

[45] A detailed analysis of the collected data for Player A can be found in the Appendix C of Chapter 5, p. 230.

[46] As almost all our participants (97.47%) rated the instructions of the experiment as understandable, we assume that subjects understood the experiment and the decisions they were asked to take, including the associated consequences.

Table 2: Descriptive data of Players B of Treatment 1 & 2

Treatment	**Treatment 1**	**Treatment 2**
n	98	100
Age (average)	22.89	22.57
Gender		
Female	54 (55.67%)	63 (63.00%)
Male	42 (43.30%)	36 (36.00%)
Non-binary	1 (1.03%)*	1 (1%)
Field of Study		
Business & Economics	52 (53.06%)	52 (52.00%)
Cultural Sciences	38 (38.78%)	35 (35.00%)
Natural Sciences	6 (6.12%)**	10 (10.00%)***
Performance in real-effort task (average payoff in tokens)		
Part 1	12.55	13.35
Part 2	4.16	4.4
Number of exclusions	12 (24.49%)	10 (20.00%)
Personal normative belief		
in support of whistleblowing	76 (77.55%)	72 (72.00%)
in support of silence	22 (22.45%)	28 (28.00%)
Normative expectation		
in support of whistleblowing	53 (54.08%)	
in support of silence	45 (45.92%)	
Empirical expectation		
in support of whistleblowing	43 (43.88%)	66 (66.00%)
in support of silence	55 (56.12%)	34 (34.00%)
Behaviour		
Whistleblowing	52 (53.06%)	59 (59.00%)
Silence	46 (46.94%)	41 (41.00%)
Social comparison orientation (average score)	3.54	3.38

*1 answer is missing
**2 answers are missing
***3 answers are missing

Test of Predictions

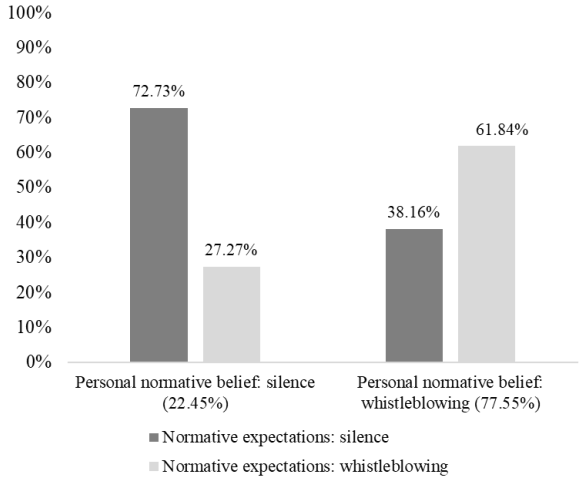
First, we examine the results of Treatment 1 to evaluate our predictions about the potential presence of norm misperceptions and the relationship between perceptions and behaviour. We then examine the results of Treatment 2 to analyse the extent to which our correcting of misperceptions, by providing information about the majority’s beliefs, has led to behavioural changes.

Misperceptions

In terms of discrepancies in beliefs, we find similar patterns in the experiment and in the survey. We observe that, out of the 76 participants who hold a personal normative belief in support of whistleblowing, 47 also hold the normative expectation (i.e., expect that others believe) that the right thing to do is to blow the whistle, and 29 that it is to remain silent. Thus, we can see that in our sample, 38.16% of those whose personal normative belief is to report wrongdoing show a discrepancy between their personal normative belief and their normative expectations. We further apply a McNemar Test to show that there is a significant difference between the rating of an individual’s own normative belief and their belief about others’ normative beliefs (McNemar’s $\chi^2 = 15.11$, $p < 0.001$). As in the survey, this indicates that there is a general discrepancy between personal normative beliefs and normative expectations.

In our sample, the vast majority of participants (77.55%) personally support whistleblowing (see Figure 2). However, 45.92% of participants indicate that they perceive that the majority of their peers in the experiment believe that the right action is to remain silent. Thus, almost half of the participants misperceive the predominant normative beliefs and therefore expect the majority of peers to believe that whistleblowing is “wrong”. We deepen the analysis by examining for which group of participants the misperceptions rather occur, whether it is amongst those with personal normative beliefs for or against whistleblowing. We observe that, out of the 76 participants who hold a personal normative belief in support of whistleblowing, 47 also believe that others share their belief, and 29 think that others oppose it. Consequently, 38.16% misperceive how the majority of others think about whistleblowing. Among the 22 participants who hold a personal normative belief in support of remaining silent, 72.72% misperceive the true majority opinion on whistleblowing, by assuming the majority also oppose it. Following the meta-analysis by Bursztyn and Yang (2022), we form a ratio between the shares of individuals who misperceive the norm and hold a personal normative belief for remaining silent (the numerator) and whistleblowing (the denominator). The value of 1.9 suggests that misperceptions are asymmetric, in that opponents of whistleblowing experience higher misperceptions than supporters. Given that we strongly observe characteristics of norm misperceptions, we find support for our first prediction.

Figure 2: Distribution of Personal Normative Beliefs and Normative Expectations among Players B in Treatment 1 (n = 98)

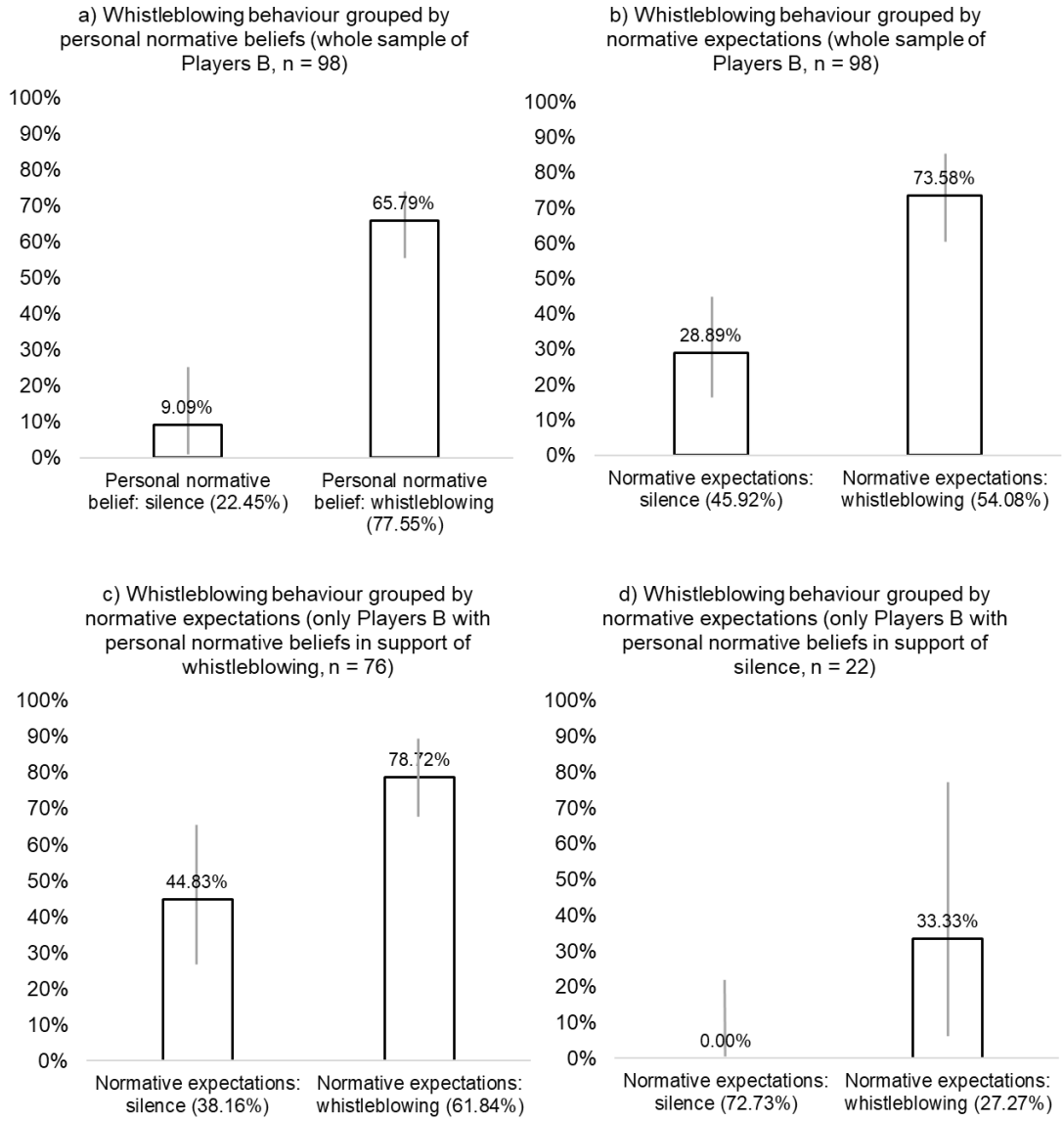


Result 1: *In the decision of whether to blow the whistle on the observed misconduct, there is a misperception about what the majority believes to be the right action. The misperceptions are asymmetric: individuals personally opposing whistleblowing exhibit a stronger misperception than those personally supporting whistleblowing.*

Personal Normative Beliefs, Normative Expectations, and Behaviour

Prediction 2 stated that holding a personal normative belief in support of reporting wrongdoing is positively related to whistleblowing behaviour compared to holding the belief that the right action is to remain silent. If we compare the share of whistleblowers and non-whistleblowers between participants holding a personal normative belief in support of versus against reporting (as displayed in Figure 3a), we observe that only two participants (9.09%) who stated their belief in remaining silent actually did blow the whistle in the experiment. By contrast, 65.79% of the participants whose personal normative belief is in favour of reporting did become whistleblowers. We test for the significance of this difference in observed behaviour by performing a one-sided Chi-Square Test (based on the hypothesized direction of the relationship) and find the relationship between personal normative beliefs and whistleblowing behaviour to be significant ($\chi^2 = 22.0214$, $p < 0.001$).

Figure 3: Personal Normative Beliefs, Normative Expectations and Whistleblowing Behaviour of Players B in Treatment 1



Note: Grey lines indicate 95% confidence intervals.

Prediction 3 focuses on the relationship between normative expectations and whistleblowing. It states that holding normative expectations in favour of reporting is positively related to whistleblowing behaviour, compared to normative expectations that the right action is to remain silent. Figure 3b shows that, among the 53 participants who think that others believe whistleblowing to be the right thing to do, 39 (73.58%) blow the whistle, while 14 (26.42%) remain silent. Among the 45 participants who expect that others hold the belief that one should remain silent, only 13 (28.89%) report the wrongdoing while the vast majority (71.11%) remains silent. A one-sided Chi-Square Test reveals that this difference is significant ($\chi^2 = 19.521$, $p < 0.001$), indicating that having

normative expectations in support of whistleblowing is significantly positively related to whistleblowing behaviour. In sum, both predictions 2 and 3 are supported.

Result 2: *Both personal normative beliefs and normative expectations are positively related to whistleblowing behaviour.*

We are particularly interested in whether normative expectations might relate to the whistleblowing behaviour of individuals whose normative belief is in support of whistleblowing. Therefore, in Prediction 4, we hypothesize that holding both a normative belief and a normative expectation in support of blowing the whistle is positively related to whistleblowing behaviour compared to normative expectations in favour of remaining silent. Figure 3c displays the distribution between normative expectations and whistleblowing behaviour for participants holding a personal normative belief pro-whistleblowing only. It shows that, of the 47 participants who are consistent in holding their personal normative belief and normative expectations in support of whistleblowing, 37 (78.72%) do blow the whistle in the experiment, while 10 (21.28%) remain silent. By contrast, out of the 29 participants whose personal normative belief in pro-reporting conflicts with their normative expectations – expecting others to disapprove of whistleblowing – 13 (44.83%) do blow the whistle, while 16 (55.17%) remain silent. According to a one-sided Chi-Square Test, normative expectations are positively related to whistleblowing behaviour, including for individuals who hold personal normative beliefs in favour of whistleblowing ($\chi^2 = 9.155$, $p = 0.001$). We therefore find support for prediction 4. We analyse the same relationship for individuals who hold a personal normative belief in favour of remaining silent. As displayed in Figure 3d, none of the 16 participants whose normative expectations match their personal normative belief in support of silence blows the whistle, while among the 6 individuals who have a normative expectation in support of whistleblowing, two (33.33%) do report the wrongdoing. Thus, for participants holding a personal normative belief against whistleblowing, normative expectations and behaviour are also significantly related ($\chi^2 = 5.867$, $p = 0.015$).

Interestingly, we find that a personal normative belief in favour of whistleblowing translates less frequently into whistleblowing behaviour than a personal normative belief against whistleblowing translates into remaining silent. Applying a McNemar Test reveals a significant difference between personal normative beliefs and actual behaviour (McNemar’s $\chi^2 = 20.57$, $p < 0.001$). Two reasons might explain the difference in

consistency between personal normative belief and behaviour. First, because of stronger misperceptions among individuals who personally oppose whistleblowing, their normative expectations are (mistakenly) more aligned with their personal views than for individuals who personally support whistleblowing. Second, whistleblowing is costly and remaining silent is not, which leads more individuals whose personal normative beliefs and normative expectations stand in conflict to choose the latter option[47].

Result 3: *For individuals whose personal normative belief is in support of (against) whistleblowing, whistleblowing (remaining silent) behaviour becomes more likely when their normative expectations are aligned with their personal belief. Personal normative beliefs translate less consistently into relevant behaviour when individuals are personally for than against whistleblowing.*

Social Information Intervention to Correct Normative Expectations

As the predominant personal normative belief in Treatment 1 is in support of whistleblowing, we inform participants in Treatment 2 that in previous sessions of the same experiment the majority of participants stated their moral support for whistleblowing. We hypothesized that providing this information prior to participants making their own decision on how to react would increase their whistleblowing behaviour. To test this conjecture, we compare the frequency of whistleblowing behaviour in Treatment 1 and Treatment 2. The results are shown in Figure 4a: in Treatment 1, 52 participants (53.06%) blew the whistle, and 46 participants (46.94%) remained silent. In Treatment 2, 59 participants (59.00%) reported the embezzlement, and 41 participants (41.00%) remained silent. Even though we observe a higher share of whistleblowers in Treatment 2 than in Treatment 1, a one-sided Chi-Square Test reveals that this difference is not significant ($\chi^2 = 0.709$, $p = 0.200$). Therefore, our prediction 5 is not supported.

However, in the next step, we only consider participants for whom the potential of being receptive to a norm message is highest, namely those who already hold a personal normative belief in support of whistleblowing. First, we check whether there is a general difference in the distribution of reported personal normative beliefs between the treatments, which might arise due to the exposure to the norm information in Treatment

[47] In the post-experimental questionnaire, we asked participants on a 1-7 Likert scale to what extent the cost of whistleblowing influenced their decision. For individuals who remained silent the cost was significantly more influential than for individuals who blew the whistle (Mann-Whitney U-Test: $z=3.335$, $p<0.001$).

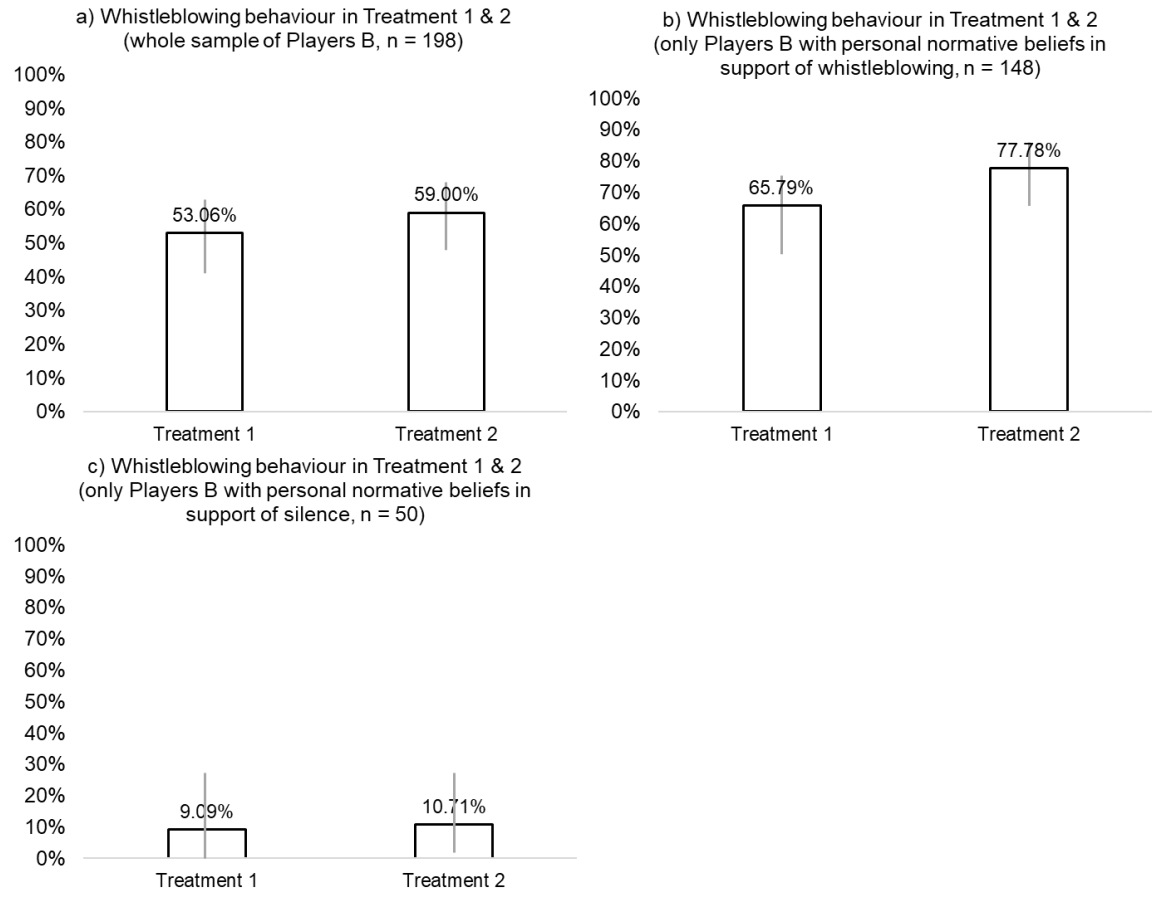
2. With 77.55% of participants expressing personal support for whistleblowing in Treatment 1, and 72.00% in Treatment 2, there is no significant difference in personal normative beliefs between the treatments (two-sided Chi-Square Test: $\chi^2 = 0.808$, $p = 0.369$), which suggests that such beliefs are relatively stable and do not vary with the exposure to the norm information. If we then compare the whistleblowing rates of these participants (see Figure 4b), we observe that in Treatment 1, 50 participants (65.79%) blow the whistle, and 26 (34.21%) remain silent, compared with 56 (77.78%) blowing the whistle in Treatment 2, and 16 (22.22%) remaining silent. This increase in whistleblowing behaviour from Treatment 1 to Treatment 2 is weakly significant when applying a one-sided Chi-Square Test ($\chi^2 = 2.614$, $p = 0.053$). Thus, for individuals who support whistleblowing, informing them that their peers share their normative belief seems to increase whistleblowing behaviour[48].

For the sake of completeness, we also test whether the social information intervention also affects the behaviour of individuals who are in support of remaining silent. Only two of the 22 subjects in Treatment 1 (9.09%) who hold that belief blow the whistle. Similarly, of the 28 subjects who express their personal support for remaining silent in Treatment 2, three (10.71%) blow the whistle (see Figure 4c). A one-sided Chi-Square Test confirms that a social information intervention does not lead to an increase of whistleblowing behaviour in that group ($\chi^2 = 0.036$, $p = 0.849$).

Result 4: *Receiving information, prior to deciding on how to respond to the observed wrongdoing, that the majority’s normative belief is in support of whistleblowing does not increase whistleblowing behaviour in general, but is effective for individuals who already support whistleblowing*

[48] We check whether there are differences in behaviour between participants who participated in the experiment online and in-person in the laboratory in the second wave of data collection. Neither in Treatment 1 ($\chi^2 = 0.2637$, $p = 0.608$) nor in Treatment 2 ($\chi^2 = 1.360$, $p = 0.244$) we find significant differences. We can therefore rule out that the environment (online vs in-person) explains the treatment effect.

Figure 4: Comparison of the Behaviour of Players B between Treatment 1 & Treatment 2



Note: Grey lines indicate 95% confidence intervals.

Regression Analyses

We corroborate our results from non-parametric testing by using a logistic regression framework (see Table 3 for the description of variables and results). First, we analyse whether personal normative beliefs and normative expectations in Treatment 1 are associated with the whistleblowing decision. In specification (1), the coefficient of personal normative belief displays a value of 0.488, which means that the likelihood of whistleblowing increases by more than 48 percentage points when an individual supports whistleblowing rather than remaining silent. Expecting others to also support whistleblowing increases the likelihood of themselves blowing the whistle by 46.1 percentage points. We see similar associations between normative expectations and

whistleblowing for subjects whose personal normative belief is in support of whistleblowing (specification 2)[49].

Second, to analyse the effect of the social information intervention on whistleblowing behaviour, we consider participants in the role of Player B from both treatments 1 and 2. The coefficients reported in specification (3) corroborate what has been reported in the previous section, i.e., that we do not find a treatment effect for the complete sample. To demonstrate that the social information intervention affects the subgroups of individuals differently, specification (4) extends the previous specification (3) by including an interaction term. As interaction effects of categorical variables in logistic regressions cannot be directly calculated and interpreted (e.g., Williams, 2012), we report the marginal effects of the social information intervention on individuals who are in support of, and on those who are opposed to whistleblowing, respectively, keeping all other variables at their mean. Conforming to the non-parametric analysis, the effect of social information intervention on individuals who personally oppose whistleblowing is almost zero (roughly 1.9% increase, i.e., non-significant). By contrast, we see an increase of 11.82 percentage points in whistleblowing among individuals who support whistleblowing. Unfortunately, while the coefficient’s sign and size corroborate the non-parametric analysis, it slightly lies above the 10% significance margin ($p=0.120$). In a final step, therefore, we specify a regression model that considers only those participants who personally support whistleblowing (specification (5)). In this model we corroborate our findings from the non-parametric analysis by showing a marginally significant coefficient with almost the same effect size as in the previous specification.

[49] Note: A specification with an interaction term between PNB and NE is not feasible with our dataset as we have no observations for the group of participants who have both PNB and NE for silence but blow the whistle. Instead, we included specification (2).

Table 3: Logistic Regression Analyses

Dependent Variable	Whistleblowing (WB) vs. Silence				
	(1) Logistic regression (only Treatment 1)	(2) Logistic regression (only Treatment 1; only Players B with PNB in support of WB)	(3) Logistic regression (both treatments)	(4) Logistic regression (both treatments	(5) Logistic regression (both treatments; only Players B with PNB in support of WB)
PNB in support of WB	0.488*** (0.101)		0.640*** (0.056)	0.638*** (0.057)	
NE in support of WB	0.461*** (0.116)	0.412*** (0.122)			
Social information intervention			0.136 (0.087)	0.125 (0.090)	0.129 (0.077)
Social information intervention:					
PNB against WB				0.019 (0.793)	
PNB in support of WB				0.118 (0.076)	
Control variables	✓	✓	✓	✓	✓
Observations	95	73	197	197	146
Pseudo-Loglikelihood	-44.961	-40.669	-100.722	-100.664	-84.110
R^2 / Pseudo-R^2	0.3167	0.1445	0.2550	0.2555	0.0389

Note: The table reports the results of a binary logistic regression, which calculated the marginal effects, with the robust standard errors shown in parentheses. The dependent variable is whether Player B blows the whistle or remains silent. Each subject constitutes one unit of observation. The reference group for personal normative belief (PNB) in support of WB is PNB in support of remaining silent, for normative expectation (NE) in support of WB it is NE in support of remaining silent, and for social information intervention it is no social information intervention. Control variables in all specifications are gender, age, field of study, and social comparison orientation. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and * respectively.

5.6 Discussion

Social norms fundamentally shape human behaviour, yet individuals frequently misperceive what others in their reference group believe or do. The misalignment between perceived and actual norms can be highly consequential when individuals adjust their actions to comply with social expectations that do not reflect reality, thereby perpetuating suboptimal social outcomes.

Our research contributes to this growing literature by conducting a comprehensive analysis that (i) identifies an economic domain with potential social norm misperceptions, (ii) examines its behavioural implications, and (iii) tests the extent to which a norm intervention that corrects for miscalibrated beliefs affects behaviour, and determines who is most responsive to such interventions.

First, we demonstrate that whistleblowing provides a compelling illustration of social norm misperceptions. A strong majority of individuals (>75%) privately support whistleblowing as an appropriate response to workplace misconduct. However, 45.92% of our participants underestimate the extent of this support among their peers, with slightly more than half correctly identifying majority support for whistleblowing. Further, we observe an asymmetry in misperceptions: the proportion of individuals who personally oppose whistleblowing and misperceive the social norm is 90% greater than the corresponding proportion of individuals who favour whistleblowing. Referring to the meta-analysis by Bursztyn and Yang (2022), these characteristics turn whistleblowing into a valuable representative case in the empirical norm misperception literature. Hence, when field data on norm misperceptions is unavailable, our experimental design may offer a controlled laboratory setting as an alternative approach to studying these phenomena. While it may be argued that the observed misperceptions could also stem from potential measurement errors (e.g., Bursztyn & Yang, 2022), a survey that we concurrently conducted on the platform Prolific with employees who have observed wrongdoing in their workplace shows a substantial indication of misperceptions in form of discrepancies between personal normative beliefs and normative expectations. Even though most employees (65%) hold a normative belief in support of whistleblowing, less than half of the subjects correctly believed that the majority also approve of whistleblowing. Twenty-five percent of subjects show the presumed discrepancy: while they hold a personal normative belief in support of whistleblowing, they expect that most of their colleagues believe that remaining silent about an observed wrongdoing is the right course of action.

Second, we show how personal normative beliefs and perceptions about the beliefs of others translate into behaviour. Both personal normative beliefs and normative expectations are positively related to whistleblowing behaviour. More generally, these findings suggest that behaviour potentially influenced by social norms can be better predicted and explained when considering both an individual’s belief about others’ views and their personal stance on the topic. Our findings complement recent experimental work

by Bašić and Verrina (2023) who examine the relevance of personal normative beliefs across various economic games. Moreover, individuals who personally support whistleblowing are more likely to blow the whistle if their normative expectations are also in support of whistleblowing than if they are against it. This behavioural pattern is in line with existing frameworks (e.g., Burks & Krupka, 2012; Krupka & Weber, 2013) that suggest that the utility of an action is higher if both personal normative beliefs and normative expectations are aligned than if they stand in conflict[50]. Consequently, what we can derive from these findings in this particular decision context is that misperceptions about the true distribution of the personal normative belief of others potentially change how a substantial share of individuals decide on taking a particular behavioural option, in our case, refraining from whistleblowing. Further, the complementarity of personal normative beliefs and normative expectations support the idea that the correcting of misperceptions could lead to behavioural change.

Third, our experimental findings reveal that providing information about the true distribution of personal normative beliefs – predominantly supporting whistleblowing – increases whistleblowing behaviour, albeit only partly. The strength of our experimental design lies in our elicitation of personal normative beliefs in the treatment implementing the social information intervention. While the intervention does not significantly increase whistleblowing behaviour across all experimental subjects, we observe a marginally significant increase among subjects whose personal normative beliefs already support whistleblowing. This finding conforms with the existing literature (Bicchieri et al., 2021; Bicchieri & Dimant, 2022a; Dimant et al., 2020; Hallsworth et al., 2017), showing that social information interventions do not guarantee behavioural change. Furthermore, these results support literature suggesting that social interventions may affect subgroups differently. As indicated in research on pluralistic ignorance where most people incorrectly believe they hold a minority opinion (Sargent & Newman, 2021), communicating information about the true distribution of personal beliefs tends to help especially those individuals who previously held a personal normative belief in support of the desired behaviour (e.g., Dannals & Miller, 2017). Additionally, as Bicchieri and Mercier (2014) suggest, achieving behavioural change is more challenging among individuals who personally oppose a behavioural option – which the social information

[50] As reflected in our survey on Prolific, where both personal normative beliefs and normative expectations are positively related to self-reported whistleblowing behaviour.

intervention tries to make attractive – than among those who support it. One counterexample is provided by Andre et al. (2024) who find that the initially opposing subgroup is affected most strongly by their intervention. The authors explain their result by assuming that this subgroup was likely surprised to learn they hold a minority view. Given that whistleblowing is a behaviour that divides opinion, with valid reasons why some people personally oppose it, this may explain our results that especially individuals who were personally against whistleblowing were not responsive to the intervention. However, as it may often be the case that through lack of appropriate theory one cannot derive predictions about how and why a social information intervention affects one subgroup more than others, future research should collect more empirical applications to identify the conditions under which differences in responsiveness can convincingly be expected, before advancing to theory building. Another finding that may interest scholars of social information interventions is the relationship between the correction of normative expectations and personal normative beliefs. Changes in our two treatment designs did not affect personal normative beliefs about whistleblowing, suggesting that majority opinions did not shift individual beliefs. However, this pattern may be context-dependent. D’Adda et al. (2020) found that personal beliefs changed when participants saw different distributions of giving behaviour in dictator games compared to the baseline with no distribution information. This suggests that personal normative beliefs can be malleable, and social information interventions that correct for misperceptions may have stronger effects on behaviour among individuals with less established beliefs compared to those holding firm beliefs that oppose majority views.

Last, our findings also contribute to the whistleblowing literature in showing that not only personal attitudes but also perceptions about how others think about it are related to whistleblowing behaviour, supporting the literature that considers whistleblowing a social phenomenon (e.g., Mir Djawadi et al., 2023; Trevino & Victor, 1992). While existing whistleblowing studies either elicited only individuals’ personal thoughts about whistleblowing (e.g., Alleyne et al., 2019) or only their normative expectations (e.g., Trongmateerut & Sweeney, 2013), our study assesses the relationship between the two in respect of whistleblowing behaviour, and how they interact with each other. Even though our main results stem from a laboratory experiment and may therefore have limited external validity, our study still carries important implications for organisations that have

adopted internal whistleblowing as an integral part of their corporate governance.[51] Misperceptions about what other employees think of whistleblowing may be a realistic outcome in organisations where the majority view on whistleblowing is either not available and/or not communicated to employees. In turn, the misperception that others disapprove of whistleblowing serves as a potential reason for why some employees may refrain from blowing the whistle. Hence, it is important to elicit the beliefs and perceptions of employees about whistleblowing even though the process of collecting this information through surveys and polls may be challenging (due to self-selection, or social desirability, etc.). By publicizing that far more employees endorse whistleblowing than is generally believed, the potential whistleblowing stigma may be attenuated and the fear of being punished by peers reduced. Additionally, the correction of misperceptions could be combined with incentives. As shown by Butler et al. (2020), financial incentives aimed at motivating employees to blow the whistle are more effective if accompanied by public social approval. As approval by work colleagues may be an even stronger motivational factor guiding behavioural conduct in the workplace, providing information about the beliefs of others alongside financial incentives may influence not only employees who personally favour whistleblowing but also those who are personally more reluctant to endorse it.

[51] Our implications are targeted at organisations that would consider that correcting the perceptions of employees about the views of peers is both desirable and potentially welfare-enhancing. Whether widespread misperceptions should always be corrected is a discussion that lies outside the scope of this study.

CHAPTER 6 | Multiple Normative Expectations and Interventions – Experimental Evidence on Whistleblowing Behaviour

Behnud Mir Djawadi, PhD Sabrina Plaß, MSc Sabrina Loer, MSc

Normative expectations – beliefs about what others perceive as appropriate – have been shown to influence behavioural choices across various domains. We examine this concept in the context of whistleblowing, where potential whistleblowers face two competing norms: promoting fairness by reporting wrongdoing versus maintaining loyalty to peers by staying silent. We investigate how normative expectations about these two mutually exclusive actions affect reporting decisions. Specifically, we test whether providing information on the majority beliefs about either the appropriateness of whistleblowing, or of staying silent, or about both behaviours together, differentially affects the whistleblowing decision. Using an incentivized experiment with UK employees on Prolific, our study yields four key findings: First, employees are more likely to report misconduct when they believe that the majority considers whistleblowing to be appropriate. Second, they are less likely to blow the whistle when they believe staying silent is deemed appropriate. Third, this effect prevails for a particularly important subgroup: among employees who believe that the majority supports whistleblowing, the reporting probability increases substantially when they simultaneously expect that staying silent is deemed inappropriate. Fourth, providing information about both normative dimensions combined or only about the inappropriateness of staying silent significantly increases whistleblowing compared to the (no information) baseline and to information about whistleblowing appropriateness alone. These findings demonstrate the importance of normative expectations about both behavioural options for accurately predicting whistleblowing behaviour, and that social information interventions are most effective when they target behaviours where appropriateness beliefs about conflicting options are disperse.

Keywords: *Whistleblowing, Normative Expectations, Social Information Intervention, Social Norms, Economic Experiment*

JEL Codes: C91, D23, D83, D9, D01

6.1 Introduction & Background

Research on social norms demonstrates that normative expectations – expectations about what others perceive as appropriate – significantly influence behavioural choice across diverse domains, from prosocial actions and charitable giving to the mitigation of discriminatory behaviour against out-group members (Agerström et al., 2016; Barr et al., 2018; Gächter et al., 2017; Krupka & Weber, 2013; Mir Djawadi et al., 2025). While organizations implement formal systems through explicit codes of conduct and compliance mechanisms, employees' ethical decisions are also subject to their perceptions of what others think and believe to be appropriate conduct (Falkenberg & Herremans, 1995; Greenberger et al., 1987). It is therefore not surprising that the development of ethical organizational culture crucially depends on members' shared beliefs and collective understanding of behaviour in situations involving ethical dilemmas (Treviño et al., 1998).

Particularly in the organizational context of internal whistleblowing[52], we believe that understanding normative expectations, and addressing them accordingly, has the potential to explain and even promote whistleblowing behaviour (see Mir Djawadi et al., 2025). We define whistleblowing as "the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action" (Miceli & Near, 1984: 4). Despite recent legislation adopted by countries and organizations and increased legal protection, internal whistleblowing is still relatively rare (Butler et al., 2020). In recent years whistleblowing has been increasingly recognised as a social phenomenon, implying that social factors play an important role in an employee's decision whether or not to report any observed wrongdoing. One reason why employees stay silent may be due to realising that "fairness and loyalty norms clash during whistleblowing decisions" (Waytz et al., 2013: 1028). Since reporting typically involves colleagues with whom one shares professional relationships (Trevino & Victor, 1992), the normative expectations regarding what is perceived by their peers to be proper conduct – reporting misconduct or maintaining loyalty towards the colleague – may be hard to guess.

[52] Unlike external whistleblowing where reports about wrongdoing go to recipients outside the organization (such as regulatory or law enforcement agencies etc.), internal whistleblowing involves reporting within the organization, usually to higher up individuals or departments that have the authority to address the problem (Dworkin & Baucus, 1998).

The co-existence of multiple, competing norms (in our case, fairness and loyalty) may have substantial behavioural implications. First, norm multiplicity can create normative disagreements that may reduce overall compliance (Dimant & Gesche, 2023). Second, the coexistence of multiple norms can provide individuals with a moral wriggle room, allowing them to selectively adhere to norms that align with their individual preferences (Merguei et al., 2022). Third, norm multiplicity may affect enforcement behaviour – when multiple legitimate normative standards exist, observers might be less inclined to punish violations of any particular norm (Panizza et al., 2023).

We therefore examine two research questions: First, how normative expectations of two mutually exclusive actions influence whistleblowing behaviour by analysing how employees' beliefs about the appropriateness of whistleblowing and of staying silent, respectively and jointly, affect reporting decisions. Second, we then test whether providing information on the majority's belief about either whistleblowing, or staying silent, or on both behaviours together, differentially affects reporting rates.

To answer our research questions, we design an incentivized economic experiment that incorporates a wrongdoing and a potential whistleblowing decision, harnessing the platform Prolific with employees from the UK. In the baseline treatment, after participants have made their decisions, we elicit normative expectations using an adapted version of the incentivized elicitation method of Bicchieri and Xiao (2009) and similar wording used by Krupka and Weber (2013). We design three additional treatments to test the different social norm interventions.

Our analysis yields four key findings. First, we find a strong positive relationship between normative expectations about whistleblowing and actual reporting behaviour: the more strongly employees believe the majority considers whistleblowing to be appropriate, the more likely they are to report misconduct. Second, we observe a clear negative relationship between the perceived appropriateness of staying silent and whistleblowing behaviour: the more strongly employees believe that silence is considered appropriate by the majority, the less likely they are to blow the whistle. Third, these effects hold also for a particular subgroup: among employees who believe whistleblowing is considered (somewhat or very) appropriate, the probability of reporting increases substantially with their expectation that staying silent is simultaneously viewed as inappropriate. This joint relationship suggests that consistent normative expectations across both behavioural dimensions has the effect of strengthening their impact on

decision-making. Finally, our treatment interventions reveal that providing information about both normative dimensions in T(WM+SM), or just on the inappropriateness of silence in T(SM), significantly increases whistleblowing, compared to the (no information) baseline and to providing information only about the appropriateness of whistleblowing in T(WM).

Our study makes important contributions to several strands of literature. First, we significantly advance the whistleblowing literature. While previous research has established that peer reactions influence whistleblowing (e.g., Taylor & Curtis, 2010; Reuben & Stephenson, 2013; Trongmateerut & Sweeney, 2013; Iwai et al., 2021; Mir Djawadi et al., 2023), our study is the first to systematically examine how normative expectations about both action (whistleblowing) and inaction (staying silent) influence behaviour. We demonstrate that both dimensions independently predict whistleblowing, and their combined effect is particularly consistent. Moreover, our finding, that more than 20% of participants believe that the majority considers both whistleblowing and silence as appropriate, challenges the prevailing assumption of pure norm polarization in whistleblowing contexts[53] (e.g., Gagnon & Perron, 2020; Olesen, 2019). This suggests that the normative landscape around whistleblowing is more nuanced than previously believed.

Second, we contribute to the social norms literature which suggests that the relationship between normative expectations and behaviour may indeed be complex. In particular, we refer to studies which showed that not only perceptions about the most frequent opinion on a given behaviour may matter, but also how normative expectations are distributed (i.e., whether a norm is considered tight or loose; Dimant, 2023) and how strong the degree of consensus regarding the norm is (D'Adda et al., 2020). In contrast to related studies that define norm uncertainty or the multiplicity of norms as the perception of different levels of appropriateness associated with different groups (Fromell et al., 2021), or the presumed discrepancy with someone else's norms or opinions about the appropriateness of outcomes (Merguei et al., 2022), we demonstrate how normative expectations about different available actions for the same decision context are related to behaviour. We show that multiple normative expectations can coexist in binary choice

[53] Polarization in the whistleblowing context involves observing the existence of two opposite profiles: one where an individual believes the majority finds whistleblowing appropriate and silence inappropriate, the other where they believe the majority finds whistleblowing inappropriate and staying silent appropriate.

settings. Further, our finding that normative expectations about both whistleblowing and silence independently and jointly affect behaviour suggests that measuring the full vector of normative expectations is crucial for understanding behavioural outcomes. Our findings extend recent studies, for example Panizza et al. (2023), who implement the Norm-Drawing Task to measure multiple normative beliefs and find that punishment behaviour of individuals holding multiple normative beliefs may be different to those with a singular normative belief.

Third, we extend the literature on social information interventions – or so-called norm-nudges[54] – by providing first systematic evidence on their effectiveness in environments with multiple competing norms. While previous research has investigated various aspects of norm-nudge design, including message content and framing (Dimant et al., 2020), and language and message inference (Kuang & Bicchieri, 2024b, 2024a), our results reveal the following novel insight: interventions are particularly effective when they target behaviours where beliefs about their appropriateness show high dispersion. In our baseline treatment T(Base), beliefs about staying silent (67% believed the majority find staying silent inappropriate) showed greater variation compared to whistleblowing (84% believed the majority find whistleblowing appropriate). This suggests that when designing single-message interventions, organizations should focus on communicating majority views about behaviours where beliefs about the appropriateness show the highest dispersion. Further, our results demonstrate that providing information about all available behavioural options, as implemented in T(WM+SM), effectively promotes whistleblowing behaviour. This finding offers organizations a practical alternative when measuring the relative dispersion of appropriateness beliefs proves to be challenging. This nuanced approach to norm-based interventions represents a significant advancement over previous one-dimensional intervention strategies, and may also be relevant for other organizational contexts where individuals face a choice between multiple actions with competing normative implications.

The paper is structured as follows: in the next section we present our predictions; in section 3 we describe our experimental set-up, procedure and treatments. Section 4 presents the results, and section 5 summarizes the main findings.

[54] Also termed (among others) norm-based interventions (Miller & Prentice, 2016), social norm messages (Bhanot, 2021) or norm cues (Cialdini et al., 1990).

6.2 Predictions

We use the utility framework by Burks and Krupka (2012) that incorporates personal normative beliefs and normative expectations, and extend it to account for norm multiplicity. An individual i's utility for choosing action a_k is characterized by:

$$u_i(a_k) = V_i\big(\pi(a_k)\big) + (N_i(a_k) - N_i(a_{-k})) + \gamma_i \left(N_g(a_k) - N_g(a_{-k})\right)$$

The function $V(.)$ represents the value the individual assigns to the monetary payoff $\pi(a_k)$ and increases in $\pi(a_k)$. The function $N_i(a_k)$ captures personal normative beliefs about action a_k, and $N_g(a_k)$ reflects the perceived normative expectations with respect to the reference group g. The more action a_k is believed to beappropriate, personally or by others, the higher are the assigned values, respectively. The novel components in our framework are the inclusion of $N_i(a_{-k})$ and $N_g(a_{-k})$, which, respectively, capture personal normative beliefs and normative expectations about the alternative action a_{-k}. This inclusion allows us to model potential norm multiplicity (for the sake of simplicity we concentrate on the two-action case, meaning that a_{-k}. represents the only alternative to action a_k). The non-negative parameter γ_i describes the individual's importance of complying with normative expectations.

Using this simple framework for the whistleblowing context, we first derive how employees' beliefs about what the majority finds (very or somewhat) appropriate/inappropriate for both actions, individually and jointly affect the selection of actions. In our setting, the two actions, whistleblowing and staying silent, are mutually exclusive behavioural choice alternatives. Note that the framework does not specify ex-ante the underlying determinants of personal normative beliefs $N_i(.)$ and the perceptions of what others view as appropriate behaviour captured by the function $N_g(.)$, or how these functions will vary across individuals or treatments. Instead, we follow Bicchieri and Xiao (2009), and employ in our experiment a two-step norm-elicitation technique to quantify personal normative beliefs $N_i(.)$, and in an incentive-compatible way normative expectations $N_g(.)$ for each action. This allows us to assess empirically the extent to which normative expectations for each action differ, and, subsequently, to examine the extent to which differences in whistleblowing behaviour are predicted by differences in the perception of how others assess the appropriateness of whistleblowing and of staying silent.

Holding constant monetary payoffs, personal normative beliefs of both actions, and majority beliefs about the alternative action a_{-k}, and referring to the framework, it is expected that normative expectations for action a_k are positively related to an individual's utility. Thus, even where multiple norms exist, individuals are more likely to choose an action that they believe others view as appropriate, and are less likely to choose that action if they believe they view it as inappropriate, ceteris paribus. For the whistleblowing context, this means that the likelihood of whistleblowing behaviour is higher if one believes that others find whistleblowing appropriate rather than inappropriate, holding all other parameters constant. As we also elicit normative expectations for staying silent, we can similarly predict what will happen to behaviour if we hold monetary payoffs, personal normative beliefs and majority beliefs about whistleblowing constant. In this case, it means that the likelihood of whistleblowing behaviour is lower if one believes that others find staying silent appropriate rather than inappropriate. These statements lead to our first predictions:

Prediction 1: Normative expectations about the appropriateness of whistleblowing are positively related to whistleblowing.

Prediction 2: Normative expectations about the appropriateness of staying silent are negatively related to whistleblowing.

Another implication from our framework is that individuals holding the normative expectation that action a_{-k} is deemed appropriate are more likely to choose a_k if they simultaneously believe that the alternative action is deemed inappropriate. This amplification occurs because consistent normative expectations across both dimensions strengthen the overall normative pull for the respective action. While normative expectations for whistleblowing have been reported in the recent empirical literature about norms and ethical behaviour (e.g., Mir Djawadi et al., 2025), to the best of our knowledge, such distribution information is not yet available for the silent option. We therefore take a conservative view and test the combined prediction only on individuals who believe that the majority finds whistleblowing appropriate. For these individuals we expect to find that the likelihood that they will blow the whistle should increase the more they also believe that others find the silent option inappropriate. Therefore, we predict:

Prediction 3: For individuals holding the normative expectation that whistleblowing is deemed appropriate (either somewhat or very appropriate), their normative expectation regarding the inappropriateness of staying silent is positively related to whistleblowing.

These three baseline predictions inform our analysis of social information interventions in settings with multiple norms. Such interventions, often termed "norm nudges", can influence behaviour by communicating what others approve or disapprove of (Bicchieri & Dimant, 2022b) – however, their effectiveness varies substantially across contexts. Previous studies document, for example, positive effects in domains like littering and plastic bag usage (Cialdini et al., 1990; De Groot et al., 2013), mixed results in tax compliance (Fellner et al., 2013; Hallsworth et al., 2017), and occasionally null effects, for example in promoting prosocial behaviour in trust games or honesty behaviour (Bicchieri et al., 2021; Dimant et al., 2020). However, social information interventions have not yet been investigated in situations where there is a choice of behavioural options, which may be perceived by some as similarly appropriate and/or inappropriate. Traditional social information interventions typically communicate either the appropriateness of a prevailing norm or the inappropriateness of a minority position, but not both at once. In contexts where there are multiple, and often conflicting, behavioural norms, this approach may be particularly problematic, because providing information about only one of the possible actions does not resolve uncertainty about the normative expectations regarding the alternative actions. Such partial information could lead to speculative inference and increased behavioural heterogeneity as individuals make different assumptions about unexpressed normative expectations. However, providing comprehensive information about both behavioural options would reduce normative uncertainty and provide clear guidance about the relative appropriateness of available actions. This leads to our final two predictions:

Prediction 4: Providing information about the majority's personal normative belief about both behavioural options (whistleblowing and staying silent) increases whistleblowing behaviour compared to communicating only the majority's personal normative belief about whistleblowing.

Prediction 5: Providing information about the majority's personal normative belief about both behavioural options (whistleblowing and staying silent) increases

whistleblowing behaviour compared to communicating only the majority's personal normative belief about staying silent.

We pre-registered these predictions (https://www.socialscienceregistry.org/trials/13850), and specified that we elicit participants' personal normative beliefs and normative expectations before they are asked to decide what behaviour to choose (and before receiving the norm messages). This sequence aimed to measure unbiased beliefs, so it was surprising that the elicitation in the baseline treatment yielded unexpectedly high whistleblowing rates (65.93%) compared to similar experiments with students (e.g., 53% in Mir Djawadi et al., 2025; 33% in Butler et al., 2020; 30% in Bartuli et al., 2016). Suspecting that this form of belief elicitation might have primed participants to behave consistently with their stated beliefs, we conducted an additional baseline treatment moving the elicitation of beliefs to after the decision. This modification significantly reduced whistleblowing rates (49.45% versus 65.93%, $\chi^2 = 5.065$, $p = 0.024$), suggesting substantial priming effects from pre-decision elicitation. Given this evidence, we implemented a post-decision elicitation procedure in our main social norm intervention treatments (all details were added to our pre-registration document). While our primary analysis focuses on these modified treatments, we report results from the initial treatments in the Appendix A of Chapter 6 (p. 233) to provide insights about the consequences of the timing of belief elicitation.

6.3 Experimental Design

Basic Experimental Set-Up

Similar to Mir Djawadi et al. (2025), our experimental design is inspired by the whistleblowing design of Bartuli et al. (2016)[55]. The basic experiment consists of two parts (for a timeline of an experimental session see Table 1). In *Part I* participants are randomly paired and collaborate on a real effort task across two rounds. For the task, team members have two minutes to individually count the number of occurrences of the digit "7" in a 300-number matrix. If their combined count falls within ±4 digits of the true value, their effort counts as successful and the team earns £1.50 per member per round.

[55] Unlike them, however, we refrain from framing the experiment in an organizational context and opt for a neutral framing. Moreover, we decided to use a neutral language in the instructions and exchanged some wording: for embezzling we used "keep part of the donation", for whistleblowing we used "reporting", and for silence we used "overlooking the behaviour of Player A".

Teams are told the outcome of their count after each round. After the second round, we randomly split the team into two, assigning one to the role of Player A, and the other to that of Player B. That round (round 3) introduces a decision situation where Player A is given the opportunity to enrich themselves fraudulently, while Player B has to decide how to respond to their team partner's wrongdoing. The roles are assigned randomly. Player A can take away part of a donation budget earmarked for the charity *Cancer Research UK*[56]. Player B can respond to Player A's embezzlement decision in two ways – either by staying silent or by blowing the whistle on their embezzlement by reporting them. If Player B stays silent, Player A increases their payoff by the amount withheld from the charity, namely £1, while Player B's earnings stay unchanged. If Player B decides to blow the whistle, the charity gets the full donation (£1.5), but both players have their payoff reduced, Player A by £1, as a punishment cost (e.g., simulating organizational sanction for the wrongdoing) and Player B by £0.50 as a whistleblowing cost (e.g., retaliation). We apply the strategy method[57] for Player B's decision, requiring them to indicate their choice before learning about Player A's actual decision. After both players have decided how to act, they are informed about the outcome and related consequences.

In Part II of the experiment, we introduce an option for social sanctions, so that when a Player B decides whether or not to blow the whistle, they may consider in their decision the potential that they may subsequently get excluded as a team player[58]. New teams are randomly matched, but comprising participants who previously held the same role (either both Players A, or both Players B). In their new teams, the roles of Player 1 and Player 2 are randomly assigned. Player 1 is informed about Player 2's choice of action in Part I of the experiment, and can decide whether they want to form a team with Player 2. This means that, the decisions that participants made in Part I potentially impacts the decision that the new team partner will make in Part II: depending on whether or not Player 1 approves of their new team mate's choice of decision in Part I, Player 2 may be excluded from the next team task. Excluding Player 2 acts as a social sanction because Player 2 is not able to earn any additional payoff, while Player 1 has to solve the task on

[56] *Cancer Research UK* (https://www.cancerresearchuk.org/) is a world leading independent cancer charity. It accounts to one of the biggest charities in terms of turnover (Rogers, 2012) and to one of the most popular charities in the UK (YouGov PLC, 2024).
[57] The strategy method is considered a common technique in experimental economics to gain a complete decision/ response structure of all participants and avoid selection issues due to non-random missing observations (Brandts & Charness, 2011).
[58] Likewise for Player A when deciding about whether or not to keep part of the donation budget.

their own, and their payoff for correctly solving the task is reduced from £1.50 to £1.20. If Player 2 is not excluded, both players solve the joint real-effort task together, as in Part I, and if they solve it correctly, they will each receive a payment of £1.50. After the task, participants are informed about the amount they earned and payment arrangements via the Prolific platform. They're also told where to find the receipt of the transfer of the whole experiment's donation to Cancer Research UK.

Elicitation of Appropriateness

For each of the team players, we asked them, first, their view on the appropriateness of the two behavioural options that had been available to them (for Player A the two options were: embezzlement and donating the full amount; for Player B: whistleblowing and staying silent) and second, how they thought the majority would view these options[59]. The elicitation took place after the two players had made their wrongdoing/whistleblowing decision. We follow the commonly used two-step elicitation procedure, also called 'opinion matching' method[60] (D'Adda et al., 2020; Görges & Nosenzo, 2020; Lane et al., 2023; Mir Djawadi et al., 2025) established by Bicchieri and Xiao (2009). We ask Player B "How appropriate do you personally believe it is to report the behaviour of Player A?", and "How appropriate do you personally believe it is to overlook the behaviour of Player A?" [61]. Participants rate these questions on a scale of four options, from "very inappropriate", via "somewhat inappropriate" and "somewhat appropriate" to "very appropriate". The form of the answer options is taken from Krupka and Weber (2013). We explain to participants that by the term "(in)appropriate" we mean that an action is (un)suitable, (un)acceptable or (not) correct for the particular circumstances.

In a second step, we elicit Player B's normative expectations about the appropriateness of both behavioural options by asking participants to assess how the majority of other participants will have answered the preceding questions. If a participant's normative

[59] In the following we describe in detail the elicitation procedure for Player B, as this is the focus of our study. We provided Player A with the same elicitation procedure regarding their behavioural options.
[60] Another commonly used method is the elicitation task by Krupka and Weber (2013). However, the Krupka-Weber method does not measure personal normative beliefs, and has a strategic component in their elicitation method that we wanted to avoid. Further, Görges & Nosenzo (2020) argue that in settings where a clear social norm might not exist (which applies to our case), the opinion matching method is the preferable elicitation technique.
[61] We apply the same elicitation procedure for Player A on the behavioural options of donating the full amount or embezzling part of it.

expectation about how the majority rates the question matches the majority's choice, they earn an additional bonus of £0.25. By telling participants that "other participants in this study are UK residents who are currently employed in an organization", we refer to a clearly defined reference group. In treatments where participants receive the social information intervention prior to their decision, eliciting normative expectations afterwards becomes redundant since these expectations have already been directly communicated through the intervention messages. However, to maintain the same bonus structures across the baseline and the three (social information) intervention treatments, we ask participants how they expect other participants to behave in the situation (empirical expectations), and incentivize their answer with a bonus of £0.25 if their assessment is correct. Participants receive the payment for correctly assessing the normative or empirical expectations via a bonus payment function on the Prolific platform.

Treatments

Treatment T(Base) serves as our baseline, measuring three key elements: personal normative beliefs regarding both whistleblowing and silence, normative expectations for both behavioural options, and actual behaviour. This design allows us to examine the presence of multiple, potentially competing normative expectations and their relationship to actual behavioural choices. The patterns observed in T(Base) are used in the design of subsequent social norm interventions aimed at promoting whistleblowing behaviour.

Based on these baseline measurements, we implement three intervention treatments: Treatment T(WM) provides participants with targeted information about whether the majority of Players B in previous sessions in T(Base) stated that reporting the behaviour of Player A is 'somewhat' or 'very' appropriate. Treatment T(SM) likewise communicates that the majority in T(Base) considered staying silent 'somewhat' or 'very' inappropriate, whereas treatment T(WM+SM) combines both messages, providing normative information about the majority's view on both of the behavioural options[62]. All three treatment interventions are delivered to participants before they are making their own decisions, to ensure that we can establish causal inferences about their effects on whistleblowing behaviour.

[62] The same information structure applied to Player A on the behavioural options of donating the money or keeping part of it.

Procedure

The experiment was programmed using oTree (Chen et al., 2016) and conducted between June 2024 (Treatment T(Base)) and November and December (for treatments T(WM), T(SM) and T(WM+SM)) via the platform Prolific Academic[63] with participants from the UK who are currently employed by an organization. We chose UK residents as participants, with the additional requirement that they are fluent in English, for two main reasons. First, we decided to recruit participants from the same country to provide them with a reference group they can more easily refer to as they share at least the same institutional and cultural background. Second, we chose the UK because Prolific originates in the UK and has the most participants there. A total of 770 subjects took part in the treatments.

At the beginning of the experiment, participants received online instructions explaining the procedure and rules of the entire experiment (Instructions in Appendix B of Chapter 6, p. 240). After completing the experiment, participants answered a questionnaire on their socio-economic demographic (including age, gender, educational qualification, job position, employment sector), questions about their donation behaviour, previous experiences of wrongdoing and whistleblowing in organizations, an excerpt from the Moral Foundations Questionnaire (Graham et al., 2011), manipulation checks and an attention check. The experimental sessions lasted 30 minutes. Participants received a flat payment of £4.00 and earned on average an additional £3.71 (exclusive of a potential bonus payment of 0.50 in the elicitation task) depending on the performance and decisions in the experiment. In total of £562.50 was donated to *Cancer Research UK* in all treatments.

[63] We deliberately chose Prolific in comparison to other crowdworking platforms. For one, participants in Prolific have been found to act with greater naïvity and less dishonesty (Peer et al., 2017). Moreover, Prolific allows to include specific requirements for participants, and for this study it is important that participants are actual employees (Peer et al., 2017). Additionally, Prolific participants show lower levels of attentional disengagement than MTurk participants, for example (Albert & Smilek, 2023). For further advantages and functionalities in comparison to other platforms we refer to Palan and Schitter (2018).

Table 1: Timeline of a Session in the Experiment

Timeline of a session.	
Stages	**Activity description**
1	Participants read a short description of the study on the Prolific platform before deciding whether to participate.
2	**The session starts** when participants enter the experiment.
3	Participants read **the instructions**, which comprehensively describe Part I and Part II of the experiment.
4	**Part I** of the experiment starts.
5	**Real effort** team task (2 rounds).
6	Participants are **informed about the success of their team task** after each round**.**
7	Participants are assigned the role of either **Player A or Player B.**
(8)	Only in Treatments T(WM), T(SM), T(WM+SM): Participants receive the **social norm intervention message.**
9	**Donation** decision (Player A); **Reporting** decision (Player B, strategy method).
10	Participants are **informed about the donation and the reporting decision** in their team, and the corresponding **consequences** are implemented.
11	**Elicitation** of personal normative beliefs and incentivized normative expectations in T(Base); Elicitation of personal normative beliefs and incentivized empirical expectations in T(WM), T(SM), T(WM+SM).
12	**Part II** of the experiment starts; teams reform, and participants become either **Player 1 or 2**.
13	Player 1 decides either to do the **team task** or the **individual task** (by excluding Player 2) based on the decision Player 2 took in Part I.
14	**Real effort** team task or individual task.
15	**Experiment ends; post-experimental questionnaire starts.**
16	**Session ends**.
17	Participants are **paid** anonymously according to their earnings in the experiment via the Prolific platform and are told where they can find a receipt of the transfer of the donation to the charity.
18	Participants in T(Base) are paid an **additional payment** if they have correctly guessed the majority's view on the appropriateness of embezzling/donating (Player A) and whistleblowing/staying silent (Player B); Participants in T(WM), T(SM), T(WM+SM) are paid an **additional payment** if they have correctly guessed the majority's actual behavioural choices of embezzling/donating (Player A) and whistleblowing/staying silent (Player B).

6.4 Results

Descriptive Data

In the following analysis, we only refer to participants in the role of Player B[64]. We had to exclude 15 participants from the analysis because they failed the attention check included in the questionnaire at the end of the experiment[65]. This results in a sample of 91 in T(Base), 90 in T(WM), 92 in T(SM) and 94 in T(WM+SM). We treat each participant as one independent unit of observation. Table 2 provides an overview of our sample across treatments, showing participants' socio-demographic characteristics, their behaviour in the experiment, their personal normative beliefs (PNB) and their normative expectations (NE) regarding whistleblowing and staying silent. Randomization achieved a good balance across treatments in terms of age (Kruskal-Wallis Test: $\chi^2(3) = 2.135$, $p = 0.5449$), gender (Chi-Square Test: $\chi^2(6) = 4.339$, $p = 0.631$), education ($\chi^2(21) = 19.287$, $p = 0.567$), and job position ($\chi^2(36) = 27.600$, $p = 0.841$). In terms of our chosen beneficiary, *Cancer Research UK*, Players B expressed their support for the charity (mean = 4.42 on a 5-point Likert scale), found it to be trustworthy (mean = 4.13 on a 5-point Likert scale), and importantly, stated that they would consider the withholding of a portion of the donation budget from *Cancer Research UK* to constitute a misconduct (mean = 3.51 on a 5-point Likert scale). This strongly suggests that participants meaningfully engaged with the ethical dimensions of the decision scenario.

As can be seen in Table 2, across all treatments, a substantial majority holds personal normative beliefs favouring whistleblowing, with 77.66% to 83.70% considering it either somewhat or very appropriate (specifically: 79.12% in T(Base), 77.78% in T(WM), 83.70% in T(SM), and 77.66% in T(WM+SM)). Correspondingly, a majority views silence as inappropriate, although with more variation across treatments (62.64% in T(Base), 56.67% in T(WM), 76.09% in T(SM), and 78.72% in T(WM+SM)). These personal beliefs align closely with normative expectations measured in T(Base), where 83.52% of participants expect others to view whistleblowing as somewhat or very appropriate, and 65.93% expect others to consider silence to be inappropriate. Hence, personal normative beliefs do match well with expectations about the views of others

[64] A detailed analysis of the data collected for Players A can be found in the Appendix C of Chapter 6 p. 243.
[65] Nearly all the participants, 99.23%, indicated that they found the instructions clear. We therefore assume that participants understood the experiment's procedure.

within our participant population. This leads us to our first key finding that emerges from the analysis of normative expectations in T(Base). As shown in Table 3, two groups represent opposing normative views that would indicate polarization: the largest group (60.44%) considers whistleblowing to be appropriate and silence inappropriate, while a small group (5.49%) holds the exact opposite view. However, the overall pattern extends beyond this polarized perspective. Nearly a quarter of participants (23.08%) believe that others view both whistleblowing and silence as appropriate, while the last group (10.99%) considers both options to be inappropriate. This distribution of beliefs challenges simplified conceptualizations of whistleblowing norms, revealing the coexistence of multiple normative standards rather than just a binary opposition between pro- and anti-whistleblowing stances[66].

Result 1: *Multiple normative expectations are observed with more than one-quarter of participants believing that others view both whistleblowing and staying silent as appropriate or both as inappropriate behaviours.*

[66] We see similar patterns when looking at personal normative beliefs (see cross-tabulation in Table A1 in the Appendix D of Chapter 6, p. 246).

Table 2: Descriptive Data of Players B across Treatments (PNB=Personal Normative Belief, NE=Normative Expectations)

	T(Base)	T(WM)	T(SM)	T(WM+SM)
Total (n)	91	90	92	94
Age (mean)	35.58	37.69	36.41	36.87
Gender				
Female	56 (61.54%)	46 (51.11%)	54 (58.70%)	50 (53.19%)
Male	34 (37.36%)	34 (37.36%)	38 (41.30%)	42 (44.68%)
Non-binary	1 (1.10%)	1 (1.10%)	-	2 (2.13%)
Education				
Student in full-time education	-	2 (2.22%)	-	1 (1.06%)
School leavers without qualification	3 (3.30%)	1 (1.11%)	1 (1.09%)	1 (1.06%)
GCSE Level	9 (9.89%)	7 (7.78%)	9 (9.78%)	10 (10.64%)
Completed apprenticeship	6 (6.59%)	5 (5.56%)	6 (6.52%)	3 (3.19%)
A-Level	9 (9.89%)	18 (20.00%)	9 (9.78%)	12 (12.77%)
Undergraduate degree	38 (41.76%)	43 (47.78%)	40 (43.48%)	45 (47.87%)
Postgraduate degree	20 (21.98%)	11 (12.22%)	24 (26.09%)	19 (20.21%)
PhD	6 (6.59%)	3 (3.33%)	3 (3.26%)	3 (3.19%)
Job position				
Upper management	6 (6.59%)	4 (4.44%)	8 (8.70%)	6 (6.38%)
Trained professional	20 (21.98%)	19 (21.11%)	11 (11.96%)	14 (14.89%)
Middle management	23 (25.27%)	22 (24.44%)	22 (23.91%)	23 (24.47%)
Skilled labourer	4 (4.40%)	5 (5.56%)	5 (5.43%)	9 (9.57%)
Junior management	6 (6.59%)	6 (6.67%)	9 (9.78%)	13 (13.83%)
Consultant	1 (1.10%)	3 (3.33%)	2 (2.17%)	-
Administrative staff	14 (15.38%)	18 (20.00%)	13 (14.13%)	13 (13.83)
Temporary employee	-	1 (1.11%)	2 (2.17%)	1 (1.06%)
Support staff	8 (8.79%)	4 (4.44%)	6 (6.52%)	6 (6.38%)
Researcher	2 (2.20%%)	2 (2.22%)	2 (2.17%)	1 (1.06%)
Student	2 (2.20%	-	2 (2.17%)	-
Self-employed/Partner	3 (3.30%)	4 (4.44%)	9 (9,78%)	6 (6.38%)
Other	2 (2.20%)	2 (2.22%)	1 (1.09%)	2 (2.13%)
Behaviour				
Silence	46 (50.55%)	44 (48.89%)	28 (30.43%)	34 (36.17%)
Whistleblowing	45 (49.45%)	46 (51.11%)	64 (69.57%)	60 (63.83%)
PNB re whistleblowing				
Very inappropriate	3 (3.30%)	1 (1.11%)	9 (9.78%)	3 (3.19%)
Somewhat inappropriate	21 (23.08%)	19 (21.11%)	6 (6.52%)	18 (19.15%)
Somewhat appropriate	32 (35.16%)	35 (38.89%)	31 (33.70%)	29 (30.85%)
Very appropriate	35 (38.46%)	35 (38.89%)	46 (50%)	44 (46.81%)
PNB re silence				
Very inappropriate	33 (36.26%)	27 (30.00%)	47 (51.09%)	33 (35.11%)
Somewhat inappropriate	24 (26.37%)	24 (26.67%)	23 (25.00%)	41 (43.62%)
Somewhat appropriate	31 (34.07%)	34 (37.78%)	18 (19.57%)	15 (15.96%)
Very appropriate	3 (3.30%)	5 (5.56%)	4 (4.35%)	5 (5.32%)
NE re whistleblowing				
Very inappropriate	3 (3.30%)	-	-	-
Somewhat inappropriate	12 (13.19%)	-	-	-
Somewhat appropriate	46 (50.55%)	-	-	-
Very appropriate	30 (32.97%)	-	-	-
NE re silence		-	-	-
Very inappropriate	26 (28.57%)	-	-	-
Somewhat inappropriate	34 (37.36%)	-	-	-
Somewhat appropriate	24 (26.37%)	-	-	-
Very appropriate	7 (7.69%)			

Table 3: Normative Expectations in T(Base)

		Normative expectation: Silence		
		Inappropriate	Appropriate	Σ
Normative expectation: Whistleblowing	Inappropriate	5 (5.49%)	10 (10.99%)	15
	Appropriate	55 (60.44%)	21 (23.08%)	76
	Σ	60	31	91

Test of Predictions

Normative Expectations and Whistleblowing Behaviour

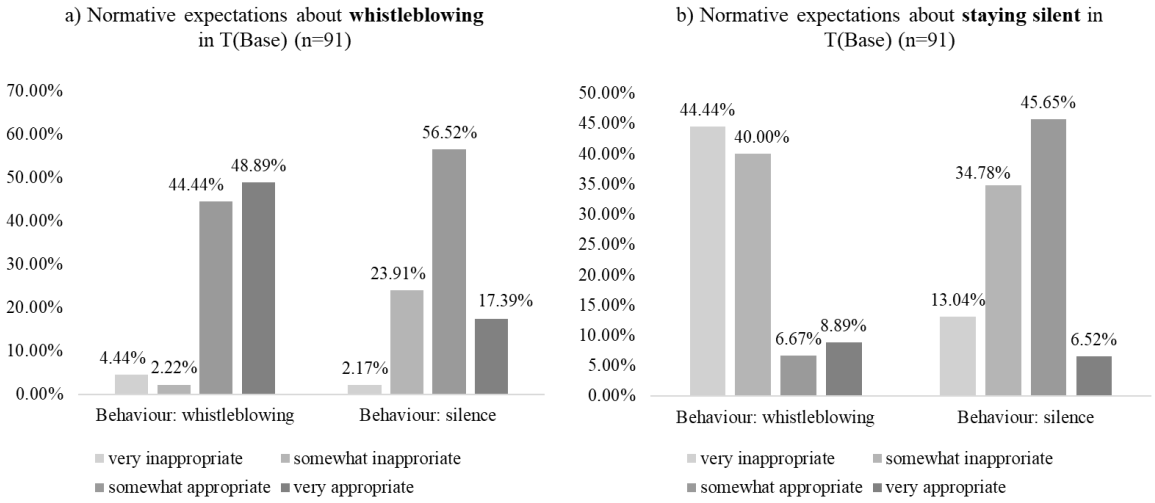
Our first analysis focuses on the relationship between normative expectations and whistleblowing behaviour in the baseline treatment T(Base), comprising 91 Players B. The whistleblowing rate in this treatment was approximately balanced, with 45 subjects (49.45%) choosing to blow the whistle and 46 (50.55%) staying silent. To quantify normative expectations, we employed a 4-point scale (1 = very inappropriate, 2 = somewhat inappropriate, 3 = somewhat appropriate, 4 = very appropriate). The baseline treatment revealed mean normative expectations of 3.15 (sd = 0.79) for whistleblowing appropriateness and 2.01 (sd = 0.75) for silence appropriateness.

Our first prediction was that the normative expectations about the appropriateness of whistleblowing are positively related to whistleblowing behaviour. The detailed distribution of the appropriateness ratings regarding whistleblowing can be seen in Figure 1a) for the respective behavioural choices. From our results, we observe that for those who blow the whistle, the mean value for the appropriateness of whistleblowing is 3.38 while it is 2.89 for those who stay silent. Comparing the two distributions, a one-sided Mann-Whitney U Test reveals that the appropriateness ratings of whistleblowing are significantly higher for whistleblowers than for subjects who stayed silent ($z = 3.449$, $p < 0.001$). Thus, we find a significantly positive relationship between the normative expectations about the appropriateness of whistleblowing, and actual whistleblowing behaviour.

Our second prediction concerned the relationship between normative expectations about the appropriateness of staying silent, and whistleblowing behaviour. For the detailed distributions of the appropriateness ratings, we refer to Figure 1b). With a value of 1.80, whistleblowers have a lower mean value for the appropriateness of staying silent

compared to subjects who did stay silent (mean value 2.46). A one-sided Mann-Whitney U Test reveals that whistleblowers' appropriateness ratings of staying silent are significantly lower than those of subjects who actually stayed silent ($z = 3.732$, $p < 0.001$). These results strongly support our prediction that normative expectations about silence are negatively related to whistleblowing behaviour.

Figure 1: Distribution of Normative Expectations and Behaviour

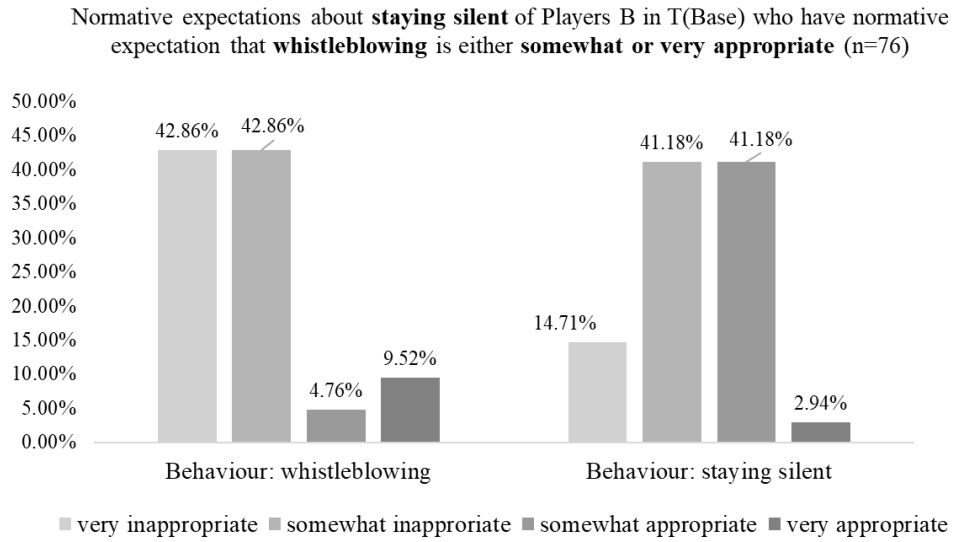


Result 2: *Normative expectations about the appropriateness of whistleblowing are positively related to whistleblowing behaviour, and normative expectations about the appropriateness of staying silent are negatively related to whistleblowing behaviour.*

In the next step, we are interested in how normative expectations about the appropriateness of whistleblowing and staying silent are jointly related to the whistleblowing decision. Therefore, we predicted that for individuals who have normative expectations that whistleblowing is appropriate (either somewhat appropriate or very appropriate), the normative expectations about the appropriateness of staying silent are negatively related to whistleblowing. The investigation of this prediction is based on a sample of 76 subjects in T(Base), whose normative expectation is that whistleblowing is either somewhat or very appropriate. In general, these subjects have a mean value for the appropriateness of staying silent of 2.04. The 42 subjects (55.26%) who blew the whistle have a mean value for the appropriateness of staying silent of 1.81, while the 34 subjects (44.74%) who stayed silent have a mean value of 2.32. The detailed distribution of appropriate ratings divided by the chosen behaviour is displayed in Figure 2. The difference in the distributions of appropriateness ratings of silence between

whistleblowers and non-whistleblowers is statistically significant (one-sided Mann-Whitney U Test: z = 2.960, p = 0.002). Thus, for individuals whose normative expectations is that whistleblowing is appropriate (either somewhat or very), the normative expectations about the appropriateness of staying silent is negatively related to whistleblowing, fully supporting Prediction 3.

Figure 2: Multiple Normative Expectations and Behaviour



For the sake of completeness, although not part of our hypothesis, we conduct the same analysis for individuals whose normative expectation is that whistleblowing is either somewhat or very inappropriate. This analysis is based on the sample of 15 Players B. Their mean value for the normative expectation of staying silent is 2.60. The 12 individuals among them who stay silent have a mean value about the appropriateness of staying silent of 2.83, while the three individuals who blow the whistle have a mean value of 1.67. This also results in a significant difference in normative expectations about the appropriateness of staying silent between whistleblowers and non-whistleblowers (one-sided Mann-Whitney U Test: z = 1.654, p = 0.049). Overall, these results strongly point to a joint relationship of the two normative expectations with actual behaviour.

Result 3: *Normative expectations about the appropriateness of whistleblowing and of staying silent are jointly related to whistleblowing behaviour.*

Social Information Interventions

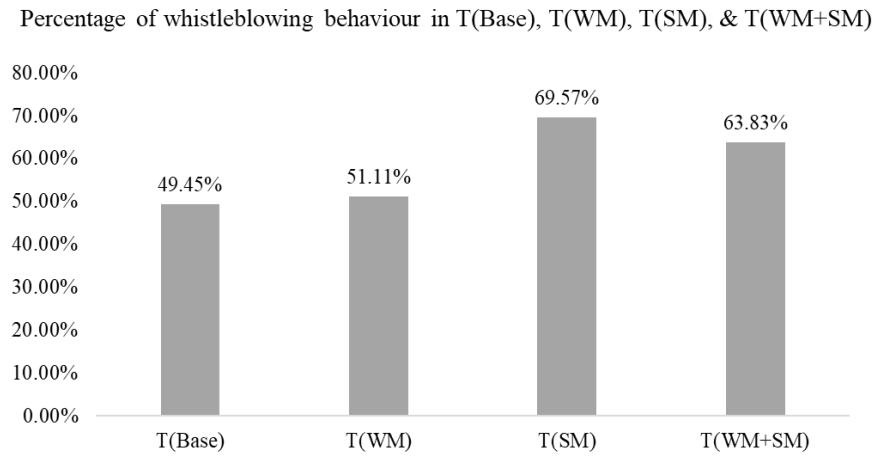
The last two predictions refer to the influence of social information interventions. More precisely, we predicted that providing information about the majority's personal normative beliefs about the two behavioural options (i.e., whistleblowing and staying silent) increases whistleblowing behaviour compared to communicating only the majority's personal normative belief about whistleblowing (Prediction 4) or about staying silent (Prediction 5).

Our analysis begins with a comparison between the baseline treatment T(Base) and the different social information treatments T(WM) and T(SM) which communicate the majority's view on the appropriateness of, respectively, either whistleblowing in T(WM) or staying silent in T(SM), and T(SM+WM), which communicate information on both. The whistleblowing rates can be found in Figure 3. We find no significant difference in actual whistleblowing rates between T(Base) and T(WM), with T(Base) showing a 49.45% whistleblowing rate and T(WM) showing a 51.11% whistleblowing rate ($\chi^2(1) = 0.050$, $p = 0.823$). This result suggests that providing social information about the majority's view on whistleblowing appropriateness alone does not significantly influence behaviour. By contrast, treatments incorporating information about the inappropriateness of staying silent show a substantial effect on whistleblowing behaviour. When comparing T(SM) with T(Base), we observe that the silence message treatment achieves a 69.57% whistleblowing rate, representing a significant 20.12 percentage point increase over the baseline treatment ($\chi^2(1) = 7.685$, $p = 0.006$), and the combined message treatment T(WM+SM) achieves a 63.83% whistleblowing rate, demonstrating a significant 14.38 percentage point increase over the baseline treatment ($\chi^2(1) = 3.895$, $p = 0.048$).

When examining the relative effectiveness of our single-message interventions, we find that the silence message treatment in T(SM) generates substantially higher whistleblowing rates compared to the whistleblowing message treatment in T(WM). Specifically, T(SM) achieves a whistleblowing rate of 69.57%, while T(WM) achieves only 51.11%. This difference of 18.46 percentage points is statistically significant ($\chi^2(1) = 6.480$, $p = 0.011$), indicating that information about the inappropriateness of staying silent is more effective at promoting whistleblowing than information about the appropriateness of whistleblowing. The comparison between the combined message treatment T(WM+SM) and the single message treatments provides additional insights.

When comparing T(WM+SM) with T(WM), we observe that the combined message treatment significantly increases whistleblowing behaviour. Specifically, of the 90 subjects in T(WM), 46 (51.11%) blow the whistle while 44 (48.89%) stay silent. In T(WM+SM), the share of whistleblowers is even higher: of the 94 subjects, 60 (63.83%) blow the whistle and 34 (36.17%) stay silent. Applying a one-sided Chi-Square Test yields a significant difference in whistleblowing behaviour between the two treatments, representing a significant improvement of 12.72 percentage points ($\chi^2(1) = 3.046$, $p = 0.041$). This finding supports Prediction 4 that adding information about the inappropriateness of silence enhances the effectiveness of the whistleblowing message. However, the relationship between T(SM) and T(WM+SM) presents a different pattern. In T(SM), of the 92 subjects, 64 (69.57%) blow the whistle while 28 (30.43%) stay silent. Comparing these shares to the 63.83% of whistleblowers in T(WM+SM) leads to a statistically insignificant difference between the treatments (one-sided Chi-Square Test: $\chi^2(1) = 0.688$, $p = 0.204$). The absence of a significant difference suggests that we have to reject Prediction 5, which stated that adding information about the appropriateness of whistleblowing to a silence message is not sufficient to give potential whistleblowers any additional motivation. This finding is particularly worth highlighting from a practical as well as a theoretical perspective, as it indicates that communicating normative expectations about the silence option alone may be sufficient to achieve the maximum effect on whistleblowing behaviour.

Figure 3: Whistleblowing Behaviour across Treatments



Interestingly, we find distinct patterns in how different types of social information affect personal normative beliefs about both whistleblowing and staying silent. Regarding

whistleblowing-related personal normative beliefs, we observe consistency across treatments. The mean personal normative belief values for whistleblowing appropriateness show no statistically significant differences in any of the treatment comparisons. As can be seen from Table 4, this consistency holds true across all treatment pairs: T(Base) vs. T(WM) (two-sided Mann Whitney U Test: $z = 0.410$, $p = 0.682$), T(Base) vs. T(SM) ($z = 1.585$, $p = 0.113$), T(Base) vs. T(WM+SM) ($z = 1.069$, $p = 0.285$), T(WM) vs. T(WM+SM) ($z = 0.706$, $p = 0.480$), T(SM) vs. T(WM+SM) ($z = 0.514$, $p = 0.607$), and T(WM) vs. T(SM) ($z = 1.289$, $p = 0.198$). This pattern suggests that participants' belief about the appropriateness of whistleblowing stays relatively stable despite exposure to different social information interventions. In contrast, the analysis of silence-related personal normative beliefs reveals significant differences. Except for the pair T(Base) and T(WM+SM), we find systematic differences between treatments that include silence-related information and those that do not. The comparison between T(Base) and T(SM) shows a significant difference in silence-related personal normative beliefs ($z = 2.120$, $p = 0.034$), indicating that exposure to information about the majority's view on staying silent affects personal beliefs about its appropriateness. Similarly, comparing T(WM) with T(WM+SM) reveals a significant difference ($z = 2.122$, $p = 0.034$), suggesting that adding silence-related information substantially influences personal normative beliefs even in the presence of whistleblowing-related information. Also, in the comparison between T(WM) and T(SM) the difference in terms of silence-related personal normative beliefs is highly significant ($z = 3.082$, $p = 0.002$). This finding may indicate that the type of social information provided (silence-related versus whistleblowing-related) has a substantial impact on personal beliefs about the appropriateness of staying silent. Our analysis reveals a consistent negative relationship between individuals' personal views on the appropriateness of staying silent, and their whistleblowing behaviour across all treatments (T(Base), T(WM), T(SM), and T(WM+SM)). This is supported by Mann-Whitney U Tests which show significant differences in the staying silent appropriateness scores between whistleblowers and non-whistleblowers (all z scores show p-values < 0.01). Thus, the increased whistleblowing rates in treatments containing silence-related information (T(SM) and T(WM+SM)) appear to be potentially driven by the interventions' success in shifting participants' personal beliefs about the inappropriateness of staying silent in response to witnessing misdemeanour inside organisations.

Result 4: *The combined communication of the majority's personal normative beliefs about both whistleblowing and staying silent in T(WM+SM) leads to significantly higher whistleblowing rates than communicating only the majority's view on whistleblowing appropriateness in T(WM). However, this approach does not outperform the treatment that solely communicates the majority's view on the inappropriateness of silence in T(SM), suggesting that silence-related normative information is of significant importance to behavioural change.*

Table 4: Two-sided Mann Whitney U Test for Personal Normative Beliefs (PNB) across Treatments

	T(Base) & T(WM)	T(Base) & T(SM)	T(Base) & T(WM+SM)	T(WM) & T(WM+SM)	T(SM) & T(WM+SM)	T(WM) & T(SM)
PNB re whistle-blowing	z = 0.410, p = 0.682	z = 1.585, p = 0.113	z = 1.069, p = 0.285	z = 0.706, p = 0.480	z = 0.514, p = 0.607	z = 1.289, p = 0.198
PNB re silence	z = 1,027, p = 0.305	z = 2.120, **p = 0.034****	z = 1,023, p = 0.306	z = 2.122, **p = 0.034****	z = 1,413, p = 0.158	z = 3.082, **p = 0.002*****

Note: Two asterisks imply the significance level of 0.05 and three asterisks indicate the significance level of 0.01.

Regression Analysis

We used a logistic regression framework to complement our findings from the non-parametric tests (see Table 5 for the description of variables and results). The first specification (1) tests how normative expectations in T(Base) are associated with the decision to blow the whistle. The coefficient of 0.184 for normative expectations about whistleblowing indicates that the likelihood of whistleblowing increases by more than 18 percentage points when an individual expects others to find whistleblowing more appropriate than staying silent. Expecting others to support silence decreases the likelihood of blowing the whistle by 17 percentage points. In specification (2), we analyse subgroup effects using simplified binary categories of the appropriateness scale: "(somewhat or very) appropriate" and "(somewhat or very) inappropriate". We report the marginal effects of the normative expectations of silence on individuals who believe others find whistleblowing (somewhat or very) inappropriate, and those who find whistleblowing (somewhat or very) appropriate, keeping all other variables at their mean. We find that for both groups in T(Base) the likelihood to blow the whistle decreases by 32.70 respectively 37.00 percentage points if others are believed to find silence (somewhat or very) appropriate. For individuals who find whistleblowing (somewhat or

very) inappropriate, this effect is not significant, probably due to the small sample size. By contrast, and corroborating our third prediction, this effect is highly significant for the group of individuals who believe others find whistleblowing (somewhat or very) appropriate. Alternatively in specification (3), when referring only to those individuals in T(Base) with normative expectations for whistleblowing being (somewhat or very) appropriate, we find that having normative expectations in favour of staying silent decreases whistleblowing behaviour by 17 percentage points, also corroborating non-parametric tests that supported our prediction about the joint relationship between normative expectations and whistleblowing behaviour.

To further analyse the effect of the treatment interventions on whistleblowing behaviour, specification (4) considers all treatments. While the social information intervention that communicates the majority view in T(WM) does not significantly affect the whistleblowing decision compared to the baseline, the social information intervention concerning silence in T(SM) significantly increases the decision by almost 20 percentage points. Providing both messages T(WM+SM) increases the likelihood of blowing the whistle by more than 13 percentage points. Specification (5) extends the findings of the previous specification in so far as it depicts the differences between providing the social information intervention about whistleblowing T(WM) versus silence T(SM) and providing both T(WM+SM). We can observe that the message about the staying silent norm in T(SM) increases the likelihood of whistleblowing by more than 20 percentage points, and providing both messages in T(WM+SM) by 14 percentage points, compared to providing a message about whistleblowing alone T(WM).

This indicates the strong influence of providing participants with the information that the majority regards staying silent to be inappropriate. This is also supported by our last specification (6), which compares the impact of the single messages (either on whistleblowing norm or on the silence norm) to each other, and similarly finds that the message about the majority belief about the inappropriateness of silence significantly increases whistleblowing compared to the majority belief about the appropriateness of whistleblowing.

Table 5: Logistic Regression Analysis

Dependent Variable	Whistleblowing (WB) vs. Silence					
	(1) Logistic regression (only T(Base))	(2) Logistic regression (only T(Base))	(3) Logistic regression (only T(Base), participants who have NE that WB is somewhat or very appropriate)	(4) Logistic regression (all treatments)	(5) Logistic regression (T(WM) versus T(SM) + T(WM+SM))	(6) Logistic regression (T(WM) versus T(SM))
NE re WB	0.184* (0.104)					
NE re Silence	-0.170** (0.083)		-0.171** (0.084)			
NE re Silence:						
NE against WB		-0.327 (0.254)				
NE for WB		-0.370*** (0.123)				
T(WM)				-0.006 (0.076)		
T(SM)				0.198*** (0.071)	0.204*** (0.072)	0.204*** (0.074)
T(WM+SM)				0.132* (0.074)	0.140* (0.075)	
Control variables	✓	✓	✓	✓	✓	✓
Observations	90	90	75	367	276	181
Pseudo-Loglikelihood	-53.300	-53.138	-47.143	-237.970	-176.135	-112.33284
R^2 / Pseudo-R^2	0.1453	0.1479	0.0874	0.0441	0.0418	0.0734

Note: The table reports the results of a binary logistic regression, which calculated the marginal effects, with the robust standard errors shown in parentheses. The dependent variable is whether Player B blows the whistle or stays silent. Each subject constitutes one unit of observation. In specification (4), the reference group for all social information interventions is T(Base). In specifications (5) and (6) the reference group for the social information interventions T(SM) and T(WM+SM) is T(WM). Control variables in all specifications are age, gender, moral foundation score fairness and moral foundation score ingroup. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and * respectively.

6.5 Discussion & Implications

Theoretical Contribution on Multiple Normative Expectations

Previous literature has explored how behaviour is shaped in situations where there is one dominating behavioural norm (e.g., Bicchieri & Xiao, 2009). However, it is not rare to find situations where more than one social norm prevails, which complicates the relationship between normative expectations and behaviour. We investigate this type of phenomenon through the lens of whistleblowing decisions. In this situation individuals are presented with two distinct behavioural options: whistleblowing and staying silent.

This context is especially suitable because normative expectations about whistleblowing may be independent from, rather than merely contrary to, expectations about staying silent. Thus, even in this binary choice situation, multiple distinct normative expectations can coexist. Our research design enables us to examine how these different, potentially conflicting normative expectations affect ethical decision-making behaviour, both independently and jointly, in an organizational context. Moreover, we investigate how behavioural changes are driven by social information interventions that communicate the majority's beliefs about the appropriateness of whistleblowing, of staying silent, or about both behavioural options simultaneously.

By extending the existing social norms frameworks (e.g., Burks & Krupka, 2012), we capture how multiple normative expectations are jointly related with each other, and how they affect the utility evaluation of actions. Our experimental findings in our baseline treatment T(Base) demonstrate that normative expectations for both behavioural options significantly affect behaviour: when individuals believe that others view whistleblowing as appropriate, whistleblowing behaviour increases, and if they believe others consider silence as appropriate, it reduces their whistleblowing. These results reveal that accurate behavioural predictions require the consideration of normative expectations for both the primary action (e.g., whistleblowing) and its alternatives (e.g., staying silent). This finding thus emphasizes the need to comprehensively examine expectations across all available behavioural options. Further, we find that normative expectations about the appropriateness of whistleblowing and staying silent are jointly related to the whistleblowing decision. In line with our extended framework, individuals who believe that others normatively support whistleblowing become less likely to blow the whistle themselves if they simultaneously believe that others normatively support the option of staying silent in that context. This finding is also supported by further statistical analysis (see Tables A2 and A3 in the Appendix D of Chapter 6, p. 246): whether considering personal normative beliefs or normative expectations, participants who view whistleblowing as appropriate and, at the same time, disapprove of staying silent, are significantly more likely to blow the whistle, compared to those who endorse both behavioural options (two-sided Chi-Square Tests: personal normative beliefs, $\chi^2(1) = 18.18$, $p < 0.01$; normative expectations $\chi^2(1) = 8.36$, $p < 0.01$). These results demonstrate that individuals holding unambiguous normative expectations make substantially different decisions from those holding multiple normative expectations. When both

behavioural options are viewed as equally appropriate, individuals tend to choose the less costly action – in our case, staying silent, which incurs no cost, versus whistleblowing, which carries associated costs. This finding extends recent work on multiple social norms. While Panizza et al. (2023) demonstrated that norm multiplicity affects how peers sanction norm violations, our findings reveal a more fundamental effect: when individuals perceive that multiple behavioural options – rather than just one dominant one – are viewed as appropriate by others, it directly influences their decision-making.

Theoretical Contribution on the Use of Social Information Interventions

Our study also advances the literature on social information interventions by examining their effectiveness in contexts with multiple normative expectations. While information about whistleblowing appropriateness alone in T(WM) did not significantly increase reporting behaviour compared to the baseline T(Base), communicating either silence inappropriateness in T(SM) or both dimensions together in T(WM+SM) effectively promoted whistleblowing. Further, compared to providing information about whistleblowing appropriateness alone in T(WM), both communicating silence inappropriateness in T(SM) and providing information about both options in T(WM+SM) significantly increased whistleblowing behaviour. The latter two treatments showed comparable whistleblowing levels. Our findings yield two key implications for social information interventions. First, these interventions appear most effective when they can resolve any lingering ambiguity in the mind of the individual about less clearly established social norms. As discussed in our hypothesis section, individuals are more likely to hold more accurate beliefs about the views of others when it comes to supporting whistleblowing than staying silent, since information about whistleblowing appropriateness may be more readily available from public sources. This is supported by our data in Table 2, which shows that, regarding normative expectations, belief distributions about staying silent are more dispersed than those about whistleblowing. Notably, within-subject comparisons in our baseline condition T(Base) reveal no significant differences between personal normative beliefs and normative expectations for either whistleblowing (Wilcoxon Test: $z = 0.388$, $p = 0.698$) or staying silent (Wilcoxon Test: $z = 0.676$, $p = 0.499$). This alignment, and the results from Table 4, suggest that our observed treatment effects were to some extent caused by updates about the actual personal beliefs of others, rather than mere compliance with perceived social norms. This interpretation aligns with D'Adda et al. (2020), who demonstrated that

exposure to information about the distribution of appropriateness ratings can reshape individuals' personal normative beliefs. Personal normative beliefs about whistleblowing stayed largely stable across all our treatments. By contrast, personal normative beliefs about staying silent proved more malleable, indicating greater initial uncertainty about the appropriateness of silence. This uncertainty may have created more opportunities for social information to actually influence personal beliefs. In environments with multiple normative expectations, our whistleblowing study reveals that single-message interventions are more effective when targeting behaviours where beliefs of appropriateness are widely dispersed. While this finding suggests a promising approach for intervention design, two important considerations emerge. First, future research needs to examine whether this principle extends beyond whistleblowing to other contexts where multiple normative expectations coexist. Second, even if this insight proves generalizable, organizations may face a practical challenge: identifying which behaviours exhibit the highest variance in beliefs of appropriateness in real-world settings can be significantly more complex than in controlled experimental conditions. Our second key implication overcomes this practical challenge by finding that providing comprehensive information about the appropriateness of both behavioural options also successfully increased whistleblowing behaviour. This result suggests a valuable alternative approach, especially when the number of different behavioural options is limited in scope and cognitive overload therefore not a major issue. Thus, when measuring the dispersion of appropriateness beliefs across options is challenging, information about all available options can be provided, which offers both practical feasibility and proven effectiveness in promoting desired behavioural changes.

Limitations

Our study has several limitations that should be considered when interpreting the results and designing future research. First, while we enhanced external validity by conducting an incentivized matching-experiment with actual employees through Prolific, this online platform approach presented certain challenges. Specifically, we experienced reduced experimenter control over the testing environment in two important dimensions. First, the 30-minute duration – relatively long for online experiments – led to participant dropouts, requiring additional data collection beyond our initial budget allocation. Second, we observed attention-related challenges that merit discussion. When asked to recall the norm message they received, a substantial proportion of participants could not

accurately retrieve this information. However, this flawed recall does not necessarily indicate a lack of comprehension or attention during the experiment, as participants may have simply struggled to remember specific message content after sustaining concentration over the extended experimental duration. Notably, we also found no behavioural differences between participants who correctly recalled the message and those who did not. Nevertheless, we acknowledge that in some cases, attention limitations related to experiment duration may have influenced participant behaviour. The second limitation concerns the methodological challenge we encountered in the timing of belief elicitation. As documented in our preregistration, we initially elicited personal normative beliefs and normative expectations before both the behavioural decision in the baseline and the norm message delivery in the treatments[67]. However, the unusually high whistleblowing rates observed in the baseline treatment indicated that the initial belief elicitation process may have unintentionally influenced participants' subsequent decisions through priming by implicitly suggesting whistleblowing as the preferred behaviour. Several studies may support our concern for the timing of belief elicitation. For example, Bicchieri (2006) suggests that asking participants to consider the appropriateness of actions before making decisions can increase norm salience and trigger social norm activation, thereby promoting norm-compliant choices. Further, experimental evidence by Gächter and Renner (2010) and Bicchieri and Chavez (2010) demonstrated that eliciting beliefs significantly influenced subsequent behaviour in public good and ultimatum games, respectively. Hence, whether our case was just an exception should be subject to further research, to systematically investigate how the timing of belief elicitation affects subsequent behaviour across different contexts. At the same time, such research would help establish more robust elicitation methods and provide more precise guidelines for experimental design, reducing reliance on trial-and-error approaches. Thus, the challenges we encountered with online platforms, maintaining attention, and the timing of belief elicitation, while important to acknowledge, also highlight the need for continued methodological refinement of experimental studies into normative beliefs and behaviour.

[67] The analysis of these initial treatments can be found in the Appendix A of Chapter 6, p. 233.

Managerial and Practical Implications

Our findings, despite their limitations, offer important implications for managers and practitioners regarding whistleblowing. The results demonstrate that expectations about others' perceptions of what constitutes appropriate and inappropriate conduct substantially impact individual decision-making, supporting emerging perspectives that deems whistleblowing a social phenomenon (e.g., Mir Djawadi et al., 2023). Further, our data challenges the notion of whistleblowing as a polarizing issue (e.g., Olesen, 2019). While we identify the two expected profiles of individuals who believe others find whistleblowing appropriate and silence inappropriate, or vice versa, more than one-quarter of participants believe that others view both whistleblowing and staying silent as appropriate or both as inappropriate behaviours. This finding reveals that knowledge about the perceived appropriateness of whistleblowing does not necessarily indicate views about the inappropriateness of silence, and vice versa. Such incomplete information can lead to inaccurate predictions about individual behaviour and the effectiveness of social information interventions that focus solely on the appropriateness of whistleblowing. These insights further suggest that organizations should gather comprehensive information about both whistleblowing and silence-related norms. The need for such a comprehensive approach is particularly evident given two key observations. First, public discourse and organizational communications typically emphasize whistleblowing-related norms, while perspectives about silence are less frequently shared or discussed. Second, organizational structure and culture can significantly influence whether a single clear norm emerges (favouring either whistleblowing or staying silent) or whether both options are viewed as equally acceptable among colleagues. Our findings suggest that addressing these normative dimensions, particularly those related to silence, can be highly beneficial. Beyond merely encouraging reporting behaviour, this approach offers practical advantages over traditional monetary incentive schemes: it can be implemented more readily and may require fewer organizational resources. Thus, by explicitly addressing the often-overlooked normative expectations about staying silent, organizations seeking to promote whistleblowing behaviour may find this communication approach a viable alternative to financial incentives.

6.6 Conclusion

We conclude that normative expectations are strongly related to whistleblowing behaviour. More specifically, we find that potential whistleblowers are i) more likely to blow the whistle if they hold normative expectations for whistleblowing, and ii) less likely if they hold normative expectations for staying silent. Moreover, these effects hold for a particularly important subgroup: among employees who believe whistleblowing is considered appropriate, reporting probability increases substantially when they expect silence to be viewed as inappropriate. In addition, providing information about both normative dimensions or the inappropriateness of silence alone significantly increases whistleblowing behaviour compared to the baseline and to information about whistleblowing appropriateness alone. Our findings provide two critical insights for understanding and promoting whistleblowing behaviour. First, accurate behavioural predictions require consideration of normative expectations for both behaviours – whistleblowing and silence – rather than focusing on either one in isolation. Second, when designing social information interventions, targeting behaviours with highly dispersed appropriateness beliefs proves particularly effective in promoting desired behavioural change.

CHAPTER 7 | Variations in the Two-Step Norm Elicitation Procedure

Sabrina Plaß, MSc

The two-step norm elicitation procedure describes a commonly used tool for measuring normative expectations in an incentivized way. This study tests some of its design features to determine whether elicited beliefs and related behaviours vary depending on i) the time of elicitation (before vs after the decision), ii) incentivizing vs not incentivizing a question about normative expectations, and iii) questioning subjects on their beliefs about the action of interest alone or combined with an alternative action. An online experiment is conducted via Prolific comprising a dictator game and the elicitation of fairness beliefs. A pretest reveals that applying role uncertainty does not alter beliefs and behaviours compared to a baseline treatment without it. Subsequently, three treatments are implemented. Contrary to previous studies, results indicate that varying the time of elicitation does not significantly alter the money-share decision. However, incentivizing the question about normative expectations significantly increases the fit with the actual majority norm. Finally, asking about a fair share and an unfair share instead of only about fair sharing does not alter personal normative beliefs or normative expectations, but it increases the empirical expectations that other dictators have provided a fair share.

Keywords: *Social Norms, Normative Expectations, Personal Normative Belief, Elicitation, Economic Experiment*

7.1 Introduction & Background

The influence of *social norms* on behaviour has sparked increasing interest in economics in the last decades. Social norms are generally described as rules about what is and is not appropriate behaviour (Görges & Nosenzo, 2020). In contrast to other-regarding preferences, we talk of a social norm when a behaviour is influenced by the dominant beliefs of a person's reference group – which can be assumed (normative expectation) or known (empirical expectation) (Bicchieri, 2006). Social norms are enforced either externally, through social punishments (e.g., avoidance, ostracism), or material sanctions (e.g., loss of property), or through internalised sanctions (e.g., feelings of guilt) (e.g., Bicchieri et al., 2018; Cialdini & Goldstein, 2004; Fehr & Fischbacher, 2004). Previous studies on social norms enhance economic theories by demonstrating that individuals do not merely act out of self-interest but also consider whether others support a particular behaviour, or behave in a particular situation (Bicchieri, 2006; Fehr & Schmidt, 2006). Hence, considerations of the utility of social approval and the potential impact on social belonging complement purely monetary considerations. The study of social norms often involves standard economic experiments, such as the dictator game (e.g., Bicchieri & Xiao, 2009; Gächter et al., 2017; Kimbrough & Vostroknutov, 2018). Experimental evidence suggests that social norms impact decision-making, especially for pro-social behaviour enacting a sense of fairness (Gächter et al., 2017; Krupka & Weber, 2013), charitable giving (Agerström et al., 2016), and in the context of honesty/lying (Abeler et al., 2019; Bicchieri, Dimant, & Sonderegger, 2023), corruption (Gneezy et al., 2019), cooperative behaviour (Reuben & Riedl, 2013), whistleblowing (Mir Djawadi et al., 2025), and discrimination against out-group members (Barr et al., 2018).

According to Bicchieri (2006, 2017), there are three types of beliefs. Personal normative beliefs describe an individual's belief about the right course of action in a given situation (first-order beliefs). Social expectations concern one's beliefs about expected behavioural norms and the corresponding behaviour of others (Bicchieri, 2017). More specifically, normative expectations describe one's perception of the majority belief of a reference group about what one *ought to do* in a certain situation, and is composed of the sum of the personal normative beliefs of others, and referred to as second-order beliefs (Bicchieri, 2017). By contrast, empirical expectations are non-normative and refer to expectations about how one perceives the majority of a reference group to behave in a

given situation. Hence, social norms represent second-order beliefs, and behavioural choices are guided by both normative and empirical expectations (Bicchieri, 2006).

The increasing importance of social norms in economics lends emphasis to the need for robust measurements to advance empirical research, and specifically to test theories of how social norms translate into behaviour. The most common approaches for measuring social norms include the non-incentivized 'belief survey' method, the incentivized 'Krupka-Weber' method (2013) for eliciting second- (or higher-)order beliefs, and the two-step elicitation method by Bicchieri and Xiao (2009), also called 'opinion matching' method, eliciting personal normative beliefs and subsequently normative expectations (Görges & Nosenzo, 2020; Lane et al., 2023). Görges and Nosenzo (2020) summarize and critically reflect upon these approaches and point out that, while neither method is superior per se, each has advantages and drawbacks, and may suit different research questions. Although the Krupka-Weber method has gained much traction lately, the two-step elicitation method by Bicchieri and Xiao (2009) continues to provide some distinct advantages over the Krupka-Weber method. First, it eliminates the latter's strategy component (no distortion of beliefs due to strategic coordination), second, it elicits both personal normative beliefs and normative expectations.[68] This is especially useful because the two-step elicitation method could reveal potential discrepancies between, on the one hand, personal normative beliefs and, on the other, normative expectations, and in so doing it may also uncover the mechanisms driving the misperception (under- or overestimating the majority's norm), or even pluralistic ignorance – the false assumption that one's own personal beliefs differ from those of the majority (Bicchieri, Dimant, & Sonderegger, 2023; Bursztyn, González, et al., 2020; Sargent & Newman, 2021).

A number of authors have applied the two-step elicitation method (e.g., Bicchieri et al., 2020, 2021, 2023; Bogliacino et al., 2024; Bursztyn, Egorov, et al., 2020; Bursztyn, González, et al., 2020; D'Adda et al., 2020; Lane et al., 2023; Mir Djawadi et al., 2025). However, while various robustness checks have been applied to the Krupka-Weber

[68] Krupka and Weber (2013)'s method elicits only normative expectations (using a Likert scale). Participants are paid a monetary reward if their rating matches the majority rating by others. However, as they are incentivized to choose the rating, they believe most others will choose, it's not even clear whether the Krupka-Weber method elicits second-order beliefs (normative expectations) or higher-order beliefs (participants' expectations about what others believe is the normative expectation) (Görges & Nosenzo, 2020).

method (e.g., Castillo et al., 2022; D'Adda et al., 2016; Fallucchi & Nosenzo, 2022; König-Kersting, 2024), only a few have been carried out on the two-step elicitation method (e.g., Aycinena et al., 2024 testing potential social desirability biases). For scholars intending to adopt the two-step elicitation method, it might become essential to pay close attention to its specific design features, for example, whether eliciting norms before or after the decision might impact decision differently.

The aim of this study is to test the robustness of variations in the two-step elicitation method, and whether (and how) these variations affect normative expectations and behaviour. We vary three different components: i) time of elicitation (before vs after the decision), ii) incentivization of normative expectations (incentivized vs non-incentivized)[69] and iii) eliciting beliefs about one vs at least two behavioural alternatives. These variations are tested in an online experiment deploying a variation of the dictator game (using role uncertainty) with UK participants recruited from the platform Prolific, and follows the general gist of Bicchieri and Xiao (2009) to elicit norms about fair sharing[70]. We test the robustness of the elicitation method in this context because previous findings have already demonstrated that fairness is a social norm (Bicchieri & Xiao, 2009). A plethora of studies using a dictator game indicates that individuals do not merely act selfishly when dividing budgets but weigh up their own payoff maximization against their social obligations (Engel, 2011). For instance, scholars indicate that dictators seem to consider what they believe is a fair share guided by personal and social rules (Bolton et al., 1998), and by the rule of reciprocity (Camerer & Thaler, 1995), while Krupka and Weber (2013) revealed that participants base their utility on taking actions that they perceive as socially appropriate in terms of a fair share.

This study contributes with four main insights to the literature on social norms and dictator games. First, we find across all treatments, that individuals believe others find a fair share less appropriate as they do. Second, behaviour does not significantly differ when eliciting normative expectations before the task compared to the elicitation after the task. Third, incentivization significantly increases the accuracy of normative expectations

[69] A critical aspect of the method is that the non-incentivization of personal normative beliefs could translate into a distortion of first-order belief, e.g., through a response bias. This, however, has been partly cancelled out by Aycinena et al. (2024), who ran several experiments and concluded that the elicitation method is not in itself prone to a social desirability bias.

[70] We follow the terms used in Bicchieri and Xiao and refer to a fair sharing for the dictator dividing the budget equally. Note that in dictator games this is often described as generosity (Engel, 2011).

(that means the correct estimation of the majority norm). Finally, asking about two behavioural alternatives – fair share and low share – increases empirical expectations of the (social norm) behaviour.

7.2 Predictions & Experimental Design

Predictions

Prediction concerning the Time of Elicitation

Based on the literature reviewed above, several predictions will be made concerning three main variations in the two-step norm elicitation procedure: i) the time of elicitation, ii) the incentivization of normative expectations, and iii) elicited beliefs on behavioural alternatives. The first of these is the time of belief elicitation, which could influence how participants behave in a task. Eliciting beliefs before the task is especially useful to investigate the impact of social information interventions on behaviour[71]. However, according to Bicchieri (2006), the salience of a norm may increase compliance with it (Bicchieri, 2006; Horne & Mollborn, 2020). Several studies have found that eliciting beliefs before the task could impact the behaviour. Evidence for this was found in experiments, for example, Mir Djawadi et al. 2015 (whistleblowing experiment), in Gächter and Renner 2010 (public goods experiment), Bicchieri and Chavez 2010 (ultimatum game). The assumption is that norm elicitation before the task could induce a framing effect and trigger the activation of social norms by making the normativity of the decision more salient. Brañas-Garza (2007) shows that behaviour in the dictator game can be prone to framing and demand effects[72], while Dreber et al. (2013) did not find support for a (far less demanding) framing effect. We propose that merely asking about personal normative beliefs and normative expectations might have a similar effect as in Brañas-Garza (2007) and nudge behaviour. Therefore, we predict as follows:

> *Prediction 1: Eliciting beliefs and expectations before the decision increases fair sharing compared to elicitation after the decision.*

[71] In experiments that investigate the influence of social information interventions, participants are usually first asked about their beliefs, and then given the information, that is, normative or descriptive messages about, for example, the majority belief from previous treatments before making a decision.
[72] In his study a sentence "Note that the receiver relies on you" influenced behaviour, which however, can be regarded as an induced social rule and a demand effect by the experimenter.

Prediction concerning the Incentivizing of the Elicitation

Studies and empirical evidence on whether or not to incentivize the elicitation of beliefs are somewhat controversial. In an ultimatum game using the Krupka-Weber method, Veselý (2015) did not find a significant difference between incentivizing and non-incentivizing the elicitation of beliefs. By contrast, incentivizing the perception of normative expectations significantly increased accuracy in a public goods game (Gächter & Renner, 2010). Moreover, it is reasoned that incentivization reduces automatic thinking (System 1 thinking) and encourages more effortful thinking (Epley & Gilovich, 2005). Accordingly, we predict that incentivizing the elicitation of normative expectations leads to participants thinking more carefully about what the most common response could be, because the incentivization increases the salience of the normative expectations. Hence, we formulate our second prediction:

> *Prediction 2: Incentivizing the elicitation of normative expectations increases the accuracy (correct estimate of the majority norm) compared to not incentivizing the elicitation.*

Predictions concerning Questions about a Fair Share and an Alternative (Low) Share

Lastly, we are interested in whether there is a difference if participants are asked about their personal normative belief and their normative expectations of the behaviour of interest only compared to being asked about at least two behaviours. In other words, we are interested in whether questions about alternative behaviours distort the perception of beliefs. As alternative share we refer to a low share, which will be investigated as the second option. Asking about the beliefs of at least two (mutually exclusive) behavioural decisions has the potential to identify whether multiple normative expectations influence behaviour. Do individuals have consistent and unambiguous beliefs and expectations in the sense that they find behaviour A being appropriate and behaviour B inappropriate? Or do individuals hold multiple inconsistent or ambiguous normative expectations by expecting two (mutually exclusive) behaviours to be similarly appropriate (or inappropriate)? Or is the behaviour even subject to polarized norms (opposing views that reinforce group divisions)?

For the context of this study, empirical evidence supports that fairness is a known and unambiguous social norm (e.g., Bicchieri & Xiao, 2009; Engel, 2011). Therefore, assessing beliefs about other behaviours should not distort the fairness norms, only

provide additional information. Hence, we do not expect any difference in personal normative belief and normative expectations of the appropriateness of fair sharing when asking about more behavioural options. In a similar vein, we do not expect that empirical expectations, the belief that other provide a fair share in our context, alter significantly either. Therefore, we assume that asking questions about the appropriateness of at least two behaviours instead of only one might affect the beliefs as follows:

Prediction 3.1: There is no difference in the appropriateness concerning personal normative beliefs of fair sharing when asking for beliefs for both behaviours, compared to eliciting personal normative beliefs about one behaviour alone.

Prediction 3.2: There is no difference in the appropriateness concerning normative expectations of fair sharing when eliciting expectations for both behaviours, compared to eliciting normative expectations about one behaviour alone.

Prediction 3.3: The number of participants having empirical expectations that others provide a fair share does not alter when eliciting normative expectations for both behaviours, compared to eliciting normative expectations concerning one behaviour alone.

Basic Experimental Set-up

The experiment mostly follows the structure of Bicchieri and Xiao (2009)'s applied dictator game. As the dictator game can be very sensitive to design elements (e.g., Brañas-Garza, 2007), we closely follow Bicchieri and Xiao's design in terms of the instructions and the increments for the money shares. However, some changes have been made, as will be explained. Instead of an on-site pen-and-paper game, UK participants were recruited via the Prolific platform to play an online version of the game. Dictators, called dividers in the instructions, have the task of dividing £2.50 between themselves and a receiver. For their share they can only choose one of the options A-G (see Figure 1). Options C and D, which give the receiver £1 or £1.25, respectively (between 40% and 50% of the total budget), are referred to as 'fair shares'. Options A and B, which give the receiver 25p (£0.25) or 50p (between 20% and 30% of the total budget), are referred to as 'low shares'[73], and options E-G (from £1.50 to £2.25) as 'high shares'. Dictators can choose to allocate the budget in 25p increments, excluding the split of £1.75 to the dictator and 75p to the receiver (exactly between a fair and a low share) and the split of 75p to the dictator and £1.75 to the receiver (exactly between a fair and a high share).

[73] Bicchieri and Xiao (2009) refer to these options as selfish shares; however, this was changed to low shares, to stay neutral in language.

Figure 1: Dividing Options

Possible options	The split
A	Dictator gets **£2.25** and Receiver gets **£0.25**
B	Dictator gets **£2.00** and Receiver gets **£0.50**
C	Dictator gets **£1.50** and Receiver gets **£1.00**
D	Dictator gets **£1.25** and Receiver gets **£1.25**
E	Dictator gets **£1.00** and Receiver gets **£1.50**
F	Dictator gets **£0.50** and Receiver gets **£2.00**
G	Dictator gets **£0.25** and Receiver gets **£2.25**

This study's design also differs by applying role uncertainty. Each participant starts off in the role of dictator and has to decide on an option[74]. After having made their decision, participants are randomly paired, with one selected to be the dictator, and the other the receiver. The dictator's chosen allocation will then be implemented.

Before proceeding with the experiment description, an explanation is due on our design's use of role uncertainty. As mentioned, scholars have previously indicated that dictator games are sensitive to design modifications, and several variations have been tested since (Camerer, 2003; Cox, 2010; Fehr & Schmidt, 2006; Kassas & Palma, 2019; Walkowitz, 2021). For instance, Heinrich & Weimann (2013) show that the dictator's behaviour is not influenced when the payoff-relevant game is chosen by the recipients compared to a random assignment. However, role uncertainty has been indicated as a potential drawback because it influences behaviour (e.g., Iriberri & Rey-Biel, 2011; Mesa-Vázquez et al., 2021; Walkowitz, 2021). Therefore, a pretest is conducted to test whether role uncertainty does alter the results from the baseline treatment. Results, more precisely depicted in Section 3.1, indicate no difference when applying role uncertainty.

[74] This is not to be confused with the strategy method (Selten, 1967), where, for a task with a first and a second mover, all possible options can be observed. By contrast, this study does not involve a second mover, but in order to gain as many observations as possible, dictator decisions are collected from all participants. Participants receive information on whether they or their teammate is in the role of the dictator after the study, which is commonly referred to as role uncertainty.

Table 1: Description of all the Stages in the Baseline Experiment T(Base), excluding Treatment Variations

Timeline of a session in T(Base).	
Stages	**Activity description**
1	Participants read a short description of the study on the Prolific platform and decide to participate.
2	Participants enter the experiment; **session starts.**
3	Participants read **general instructions.** Instructions provide complete information about of the experiment.
4	**Experiment** starts.
5	**Dictator** decision (all participants).
6	Elicitation of **personal** normative beliefs and **normative expectations** (incentivized).
7	Elicitation of **empirical expectations**.
8	**Experiment ends; Survey starts.**
9	**Session ends**.
10	Participants are **paid** the fixed payment (£1) anonymously via the Prolific platform.
11	Participants are randomly **matched**; randomly one is chosen to be dictator and the other receiver; the decision of the dictator will be implemented.
12	Participants are paid an additional **bonus payment** according to the decision of the dictator and if they are correct in the elicitation of the appropriateness.

After subjects have decided on the budget share, the experiment elicits participants' personal normative beliefs and normative expectations (the most common response) regarding the fair share (Option C or D).[75] Participants are asked to select whether they personally believe providing a fair share is very inappropriate, somewhat inappropriate, somewhat appropriate or very appropriate. To elicit their normative expectations, participants are asked which answer they think most of the other participants did choose in the preceding question, presenting them with the same four response options (very inappropriate – very appropriate). Normative expectations are incentivized in that

[75] This is a deviation from Bicchieri and Xiao (2009), who ask personal normative beliefs as a dummy variable ("Do you think you should make a fair offer? / Do you think that dividers should split the money approximately equally (choose C or D)?) and elicit a concrete number for normative expectations ("How many dividers in this room do you think answered 'Yes' to question (d)?"). Instead, to account for more nuances, this study employs a Likert scale for participants to indicate the appropriateness of the actions. This adaption was needed because giving a concrete number in the online experiment was not feasible due to potential dropouts. The formulation of the questions is similar to that used by Krupka and Weber (2013), but the specification "socially" in front of 'appropriate' has been dropped.

participants will receive an additional bonus of 25p if their answer to the question matches the correct majority choice. On the next screen, participants are asked about empirical expectations, more specifically, whether they expect others to provide a fair share (yes/no). Bicchieri (2006) emphasizes the importance of the reference group in the context of social norms. Hence, our participants are informed that all the other participants in this study are UK residents. Additionally, it is explained that the term "inappropriate" means an unacceptable, unsuitable or incorrect action in that situation. The term "appropriate" indicates an acceptable, suitable or correct action in that situation.

Lastly, participants are asked to complete a questionnaire about their demographics, control variables, fairness items of the Moral Foundation Questionnaire (Graham et al., 2011), and two attention check questions. In a baseline treatment T(Base), the experiment is conducted as explained above and depicted in Table 1. A pretest is conducted to check whether role uncertainty distorts beliefs and behaviours compared to T(Base).

Treatment Variations

Three treatments will be conducted implementing different variations of the elicitation procedure (see Figure 2). In treatment T(Before), the elicitation time varies. Everything else is held constant, but the elicitation of beliefs and expectations is conducted before the decision, while in T(Base), the elicitation takes place after the decision. In treatment T(NoIncentive), the additional incentivization for correctly stating the majority norms is dropped, whereas in T(Base), participants are externally incentivized to carefully consider the question to receive a bonus payment (if their estimation equals the most common response). All other aspects are held constant. In treatment T(Questions), in addition to being asked about the appropriateness of a fair share, participants are also asked about their personal normative beliefs and the normative expectation (incentivized) of a low share (Option A or B)[76].

[76] The question about personal normative beliefs of a low share states: "How appropriate do you personally believe it is to make a low offer to the Receiver (Option A or B)?", with the potential answers ranging from very inappropriate, somewhat inappropriate, somewhat appropriate to very appropriate. The question on normative expectations asks: "Which answer do you think the majority of participants chose in the preceding question? If your answer matches the actual answer of the majority, you will earn an additional 25p."

Figure 2: Treatment Variations / Manipulations

	T(Base)	T(Before)	T(NoIncentives)	T(Questions)
Time of elicitation	after	before	after	after
Incentivized elicitation	yes	yes	no	yes
Questions about a low and a fair share	fair share	fair share	fair share	both

An a-priori sample size calculation, assuming a small to medium-sized effect (Cohen's d: 0.35; probability level: 0.05; statistical power level: 0.8), reveals that 102 observations per treatment are needed. Each participant is assigned to only one of the treatments.

Procedure

The experiment was programmed using oTree (Chen et al., 2016) and conducted in December 2024 via the platform Prolific Academic[77] with participants from the UK. There are three main reasons why UK residents were chosen as participants, one of which was the requirement to be fluent in English (criteria set in Prolific). Second, recruiting participants from the same country provides them with a reference group they can more easily refer to as they share at least broadly the same institutional and cultural background. Third, Prolific originates from the UK and has the most participants there. The study received prior ethical approval, and the study has been pre-registered[78]. Participants received the instructions at the beginning of the experiment[79].

7.3 Results

Pretest

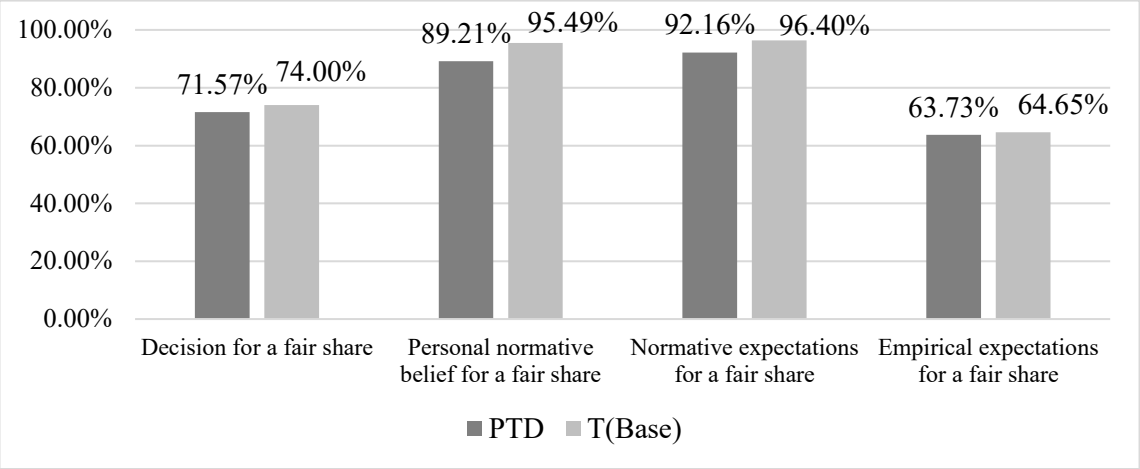
Before conducting the treatments, we run a pretest to investigate whether role uncertainty influences behaviour and beliefs in the studies' experimental design. The

[77] We deliberately chose Prolific in comparison to other crowdworking platforms. Participants on Prolific have been found to be more naïve and less dishonest than, say MTurk (Peer et al., 2017); Prolific allows the inclusion of specific requirements for participants (Peer et al., 2017), and Prolific participants show lower levels of attentional disengagement than MTurk participants (Albert & Smilek, 2023). For further advantages and functionalities compared to other platforms refer to Palan and Schitter (2018).
[78] For the ethical approval from the GfeW see: https://gfew.de/ethik/Bm3XqJ6f and for the peer-reviewed re-registration of the study see Social Science Registry: https://www.socialscienceregistry.org/trials/14794.
[79] For the instructions see Appendix B of Chapter 7, p. 249.

pretest and its findings are briefly described before reporting the main results. The pretest aims to investigate whether the results differ from those using the standard procedure in T(Base). In T(Base) all participants perform the task as dictators and, to implement role uncertainty, they are randomly assigned to either the role of *dictator* (called divider in experiment*)* or *receiver* only at the end. By contrast, the pretest performs a dictator game without using role uncertainty, where participants are informed of their role of dictators or receivers in the study beforehand (divided for analysis into PTD=Pretest Dictator and PTR=Pretest Receiver). The share provided by the dictators in PTD, and the norms elicited from them, are then compared to T(Base).

In total, 201 subjects participated in the pretest with 102 as dictator (PTD) and 99 as receiver (PTR, some failed attention checks or did not finish), and 111 in T(Base). The average amount given to receivers is £0.95 in PTD and £1.00 in T(Base). Additionally, the mean value of elicited beliefs is similar, with personal normative beliefs on average measuring 3.42 in PTD and 3.55 in T(Base), and normative expectations 3.27 in PTD and 3.35 in T(Base). Figure 3 depicts the percentage differences between dictators of PTD and T(Base) (for a more detailed analysis, see Table A1 in Appendix A of Chapter 7, p. 247). Results reveal no significant difference in personal normative beliefs about a fair share (Chi-Square Test: $\chi 2$ (3) = 3.1156, p = 0.374), no significant difference between normative expectations ($\chi 2$ (3) = 1.9994, p = 0.573) and no difference in empirical expectations ($\chi 2$ (1) = 0.0100, p = 0.920). Moreover, the decision for a fair share did not differ significantly (two-sided Mann-Whitney U Test: z = -0.561, p = 0.5745). Therefore, the subsequent treatments use role uncertainty in the dictator game, where each participant is one observation unit.

Figure 3: Descriptive Data of the Pretest (Percentage Differences between PTD and T(Base)



Descriptive Data

Across all treatments, 28 participants who did not finish the study or failed one of the attention checks were excluded. The final sample totals 438 subjects, of which 111 are in (T(Base), 109 in T(Before), 110 in T(NoIncentive) and 108 in T(Questions). More than 99% of participants found the instruction comprehensible. Participants received a fixed payment of £1 and on average an additional £1.38 as a bonus payment (on average in T(Base): £1.35, in T(Before): £1.37, in T(NoIncentive): £1.33, and in T(Questions): £1.48)[80].

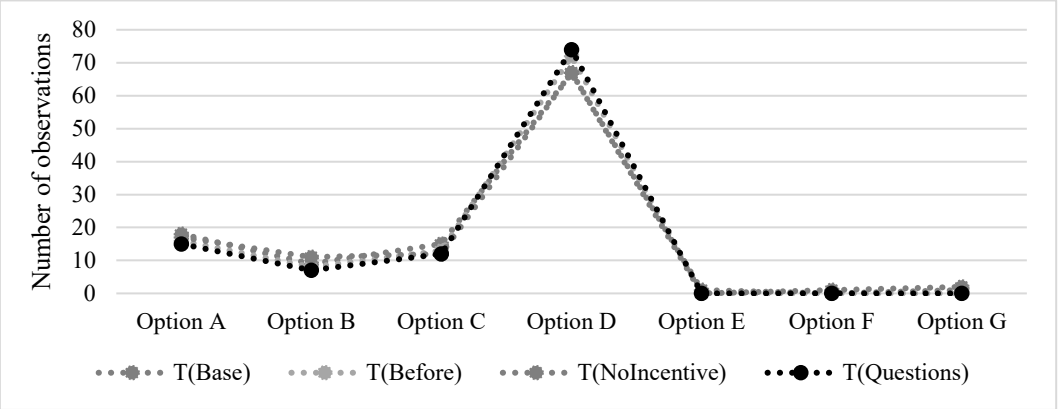
There is no significant difference in the treatment compositions concerning age (Kruskal-Wallis Test: χ^2 (3) = 2.255, p = 0.5212), gender (Chi-Square Test: χ^2(6) = 7.5728, p = 0.271), education (χ^2 (21) = 13.0407, p = 0.907), whether or not participants are in employment (χ^2 (6) = 4.4317, p = 0.618), or which role they have at work (χ^2 (39) = 34.8594, p = 0.659). Moreover, there is no significant difference between the share in the decision task (χ^2 (18) = 11.8959, p = 0.853). The majority in each treatment decided on Option D, that is, one defined as a fair share (see Table 2, Figure 5 and Table A1 in the Appendix, p. 247).

The majority of dictators across all treatments share the budget equally (Figure 4). More precisely, dictators granted receivers on average about 40% of the money (mean in T(Base): £1.00, in T(Before): £1.02, in T(NoIncentive): £1.02, and in T(Questions):

[80] This aligns with the average bonus payoff in the Pretest, which amounts to £1.36.

£1.03).[81] These results are in line with previous studies inferring participants are not profit-maximizing but hold norms for fairness and generosity. In a meta-analysis, Engel (2011) finds that across numerous dictator games (616 treatments from 129 studies) dictators provide on average 28%. Several other studies show that the majority provides an equal share in the dictator game and expects this to be normatively the right thing to do (e.g., Andreoni & Bernheim, 2009; Kimbrough & Vostroknutov, 2016, 2018; Krupka & Weber, 2013).

Figure 4: Dictators Behaviour across Treatments



Note: Option A and Option B were classified as a low share, Options C and D as fair share, and Options E to G as high shares.

On average, 90% of the participants hold the personal normative belief that a fair share is (somewhat or very) appropriate (Table 2, Figure 5). This aligns with normative expectations, where on average 90% expect a fair share to be appropriate. Interestingly, even though both normative expectations and personal normative beliefs are regarded as (somewhat or very) appropriate by the majority, the degree of appropriateness is distributed in the opposite way: the majority expect others to find a fair share only *somewhat* appropriate (52.22%) but have personal normative beliefs of a fair share being *very* appropriate (59.26%). By contrast, 38.88% expect the normative expectation of a fair share to be *very* appropriate, while 33.89% personally believe a fair share is *somewhat* appropriate.

[81] We cannot directly compare these results with Bicchieri and Xiao (2009) because they do not have a neutral or baseline treatment, but all treatments contain a specific norm message and are compared with one another.

Table 2: Descriptive Data of Dictators

Behaviour	**Low share**	**Fair share**	**High share**	**Share = majority norm** (D in all treatments)
T(Base)	24.33%	73.87%	1.8%	60.36%
T(Before)	24.02%	77.98%	/	60.91%
T(NoIncentive)	25.45%	71.82%	2.73%	68.51%
T(Questions)	20.37%	79.63%	/	62.41%

Empirical expectations	**Fair share (yes)**	**Fair share (no)**
T(Base)	63.06%	36.94%
T(Before)	69.72%	30.28%
T(NoIncentive)	62.73%	37.27%
T(Questions)	74.07%	25.93%

Personal normative beliefs about a fair share	**Very inappropriate**	**Somewhat inappropriate**	**Somewhat appropriate**	**Very appropriate**
T(Base)	2.70%	1.80%	33.33%	62.16%
T(Before)	0.92%	1.83%	36.70%	60.55%
T(NoIncentive)	1.82%	10%	28.18%	60%
T(Questions)	2.78%	1.85%	41.67%	53.70%

Normative expectations about a fair share	**Very inappropriate**	**Somewhat inappropriate**	**Somewhat appropriate**	**Very appropriate**
T(Base)	1.80%	1.80%	55.86%	40.54%
T(Before)	1.83%	4.59%	43.12%	50.46%
T(NoIncentive)	1.82%	15.45%	54.55%	28.18%
T(Questions)	1.85%	7.41%	54.55%	35.19%

The according mean values reinforce that normative expectations are lower than personal normative beliefs across the treatments. T(Base) revealed a mean value of personal normative beliefs of 3.55 (sd = 0.06) for the appropriateness of a fair share, and values in other treatments are similar (3.57 (sd = 0.06) in T(Before), 3.46 (sd = 0.07) in T(NoIncentive), and 3.46 (sd = 0.07) in T(Questions)). Additionally, the mean value for normative expectations of a fair share is similarly distributed across treatments with a value of 3.35 (sd = 0.06) in T(Base), 3.42 (sd = 0.07) in T(Before), 3.09 (sd = 0.07) in T(NoIncentive), and in T(Questions) it is 3.24 (sd = 0.08). Additional analysis supports the finding that personal normative beliefs significantly differ from normative expectations in all treatments. As mentioned, while the majority of participants find a fair share very appropriate, the majority expects that others hold the belief that a fair share is only somewhat appropriate (Wilcoxon signed-rank test: T(Base): z = 3.476, p = 0.005;

T(Before): $z = 2.595$, $p = 0.0095$; T(NoIncentive): $z = 4.389$, $p < 0.00$; T(Questions): $z = 3.124$, $p = 0.0018$; PTD: $z = 2.664$, $p = 0.0077$). Empirical expectations are mostly similarly distributed among treatments, where two-thirds (on average 66.67%) of the participants expect others to provide a fair share (see Figure 5a-c for percentages across treatments).

Result 1: *While the majority find a fair share very appropriate, the majority expect others to find a fair share only somewhat appropriate. However, there is no discrepancy between personal normative beliefs and normative expectations of whether providing a fair share is appropriate or not.*

Figure 5: Distribution (Percentage) Regarding a Fair Share across Treatments

Test of Predictions

Results regarding Prediction 1 – Time of Elicitation

Concerning the time of norm elicitation, we find no significant difference between eliciting personal normative beliefs on the appropriateness of providing a fair share (Option C or D) before in T(Before) or after the decision in T(Base). Applying a one-sided Mann-Whitney U Test reveals that the appropriateness ratings of a fair share are not significantly different ($z = 0.127$, $p = 0.44955$). Our results also reveal that normative expectations measured before and after the decision do not differ significantly (one-sided

Mann-Whitney U Test: z = -1.148, p = 0.12545). However, slightly significantly more participants estimated the correct normative expectation (that a fair share is very appropriate) when eliciting normative expectations before in T(Before) than after the decision (one-sided Chi-Square Test: $\chi^2(1) = 2.182$, p = 0.07). This difference in accuracy may be due to participants having a more positive view of the normativity of others before than after having themselves made the decision.

In respect of actual behaviour, results indicate that decisions to provide a fair share are not significantly different when beliefs are elicited before and after the task, leading us to reject H1 (one-sided Chi-Square Test: $\chi^2(1) = 0.5075$, p = 0.238). Empirical expectations do not differ either (two-sided Chi-Square Test $\chi^2(1) = 1.0933$, p = 0.296).

Result 2: *Behaviour does not significantly differ when eliciting personal normative beliefs and normative expectations before the task compared to after the task, leading us to reject Prediction 1.*

Results regarding Prediction 2 – Incentivizing Elicitation

Normative expectations for a fair share are considered significantly more appropriate when they are incentivized in T(Base) compared to when they are not in T(NoIncentive) (one-sided Mann-Whitney U Test: z = 2.860 p = 0.0021). Referring to the second Prediction, applying a one-sided Chi-Square Test reveals that the incentivization of normative expectations has a significant impact on accuracy ($\chi^2(1) = 3.7400$, p = 0.0265). The share of correct fits (normative expectations matching the most common actual response) is higher in T(Base) with incentivization than in T(NoIncentive) without an incentive – supporting Prediction 2. Moreover, we find no significant differences between the two treatments regarding personal normative beliefs (two-sided Mann-Whitney U Test: z = 0.685, p = 4935), behaviour (z = -0.237, p = 0.1824), and empirical expectations (two-sided Chi-Square Test $\chi^2(1) = 0.0027$, p = 0.959).

Result 3: *Incentivization significantly increases the accuracy of normative expectations, supporting Prediction 2.*

Results regarding Predictions 3.1-3.3 – Questions about a Fair and Low Share

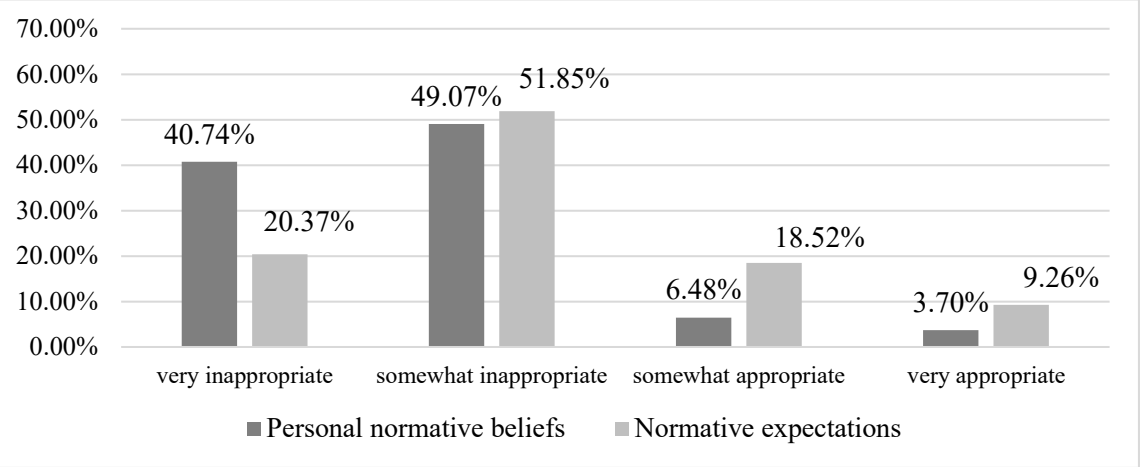
When asking participants about their belief about not just one behavioural action (a fair share), but at least two behavioural alternatives (fair share and low share) both personal normative beliefs and normative expectations about the appropriateness of a fair

sharing decrease. Put differently, when participants were asked to consider their own beliefs not only about whether it's right to provide a fair share, but also whether it is valid to provide a low share, they were less likely to believe in the appropriateness of providing a fair share, or to believe that others shared that view. However, the decrease is neither significant either for personal normative beliefs in Prediction 3.1 (two-sided Mann-Whitney U Test: $z = 1.198$, $p = 0.2311$) nor for normative expectations in Prediction 3.2 (two-sided Mann-Whitney U Test: $z = 1.226$, $p = 0.2204$). Regarding empirical expectations, our results indicate that significantly fewer participants in T(Base) expect others to provide a fair share, compared to those in T(Questions) (one-sided Chi-Square Test (1) = 3.0754, $p = 0.0395$), leading us to reject Prediction 3.3. Asking personal normative beliefs and normative expectations about both alternatives might have increased the salience of a fair share. Afterall, most participants indicated a personal normative belief and normative expectations of a low share as inappropriate, which might have increased the expectation that others provided a fair share.

Results 4: *Asking about at least two behavioural alternatives instead of only one (the main option of interest) does not significantly influence personal normative beliefs and normative expectations, providing support for Predictions 3.1 and 3.2. But it increases empirical expectations about the behaviour of interest, leading us to reject Prediction 3.3.*

In the next section, we further analyse participants answers regarding a low share. Concerning answers about a low share, most participants hold personal normative beliefs (89.91%) and normative expectations (72.22%) for a low share to be (somewhat or very) inappropriate. In contrast, only 10.18% of the subjects stated that they find providing a low share (somewhat or very) appropriate, while more than twice as many (27.78%) expect others to believe a low share to be appropriate. About half of the participants held personal normative beliefs and normative expectations of a low share to be somewhat inappropriate (Figure 6). Interestingly, only half as many believe a low share to be very inappropriate (40.74%) and expect others to share this belief (20.37%). By contrast, while only about 10% personally believe a low share to be either somewhat or very appropriate, their number nearly triples for normative expectations (27.78%) (Figure 6). The mean value for normative expectations of a low share in T(Questions) is 2.17 (sd = 0.08), and for personal normative beliefs it is 1.73 (sd = 0.07) – a significant difference from normative expectations (Wilcoxon signed-rank test: $z = -4.499$, $p < 0.00$).

Figure 6: Percentage Distribution of a Low Share in T(Questions)



After all, asking about two behavioural alternatives has the advantage of providing data for additional analyses. For instance, it allows to investigate whether participants have multiple normative expectations (e.g., expecting that others find two opposing actions similarly appropriate or inappropriate), whether the behaviour is subject to polarized norms (e.g., some individuals find one action appropriate and the other inappropriate, while for others the exact opposite is true), or whether an unambiguous norm prevails. In this study, combining both normative expectations, we find that two-thirds (66.67%) of the participants expect a fair share to be appropriate and a low share to be inappropriate (Table 4), indicating a consistent expectation in favour of a fair share for the majority. Still, one-quarter of participants expect both to be appropriate, indicating the presence of multiple normative expectations. This may also result from the fact that normative expectations for a low share were more widely distributed than those for a fair share. Expecting both behaviours to be appropriate might create cognitive dissonance and uncertainty about how to behave. In this case, it might be worth investigating whether providing information about the true majority belief (a low share being inappropriate) might resolve this ambiguity, thereby reducing cognitive dissonance and leading to an even greater increase in empirical expectations and behaviours favouring a fair share.

Results 5: *In line with the majority norm, the majority has a normative expectation that a fair share is appropriate and a low share inappropriate. However, one-quarter finds both actions appropriate, which does not reflect the majority norm, and which might be resolved by providing information on the majority norm concerning a low share.*

Table 3: Multiple Normative Expectations

		Normative Expectation: Fair Share		
		Inappropriate	Appropriate	Σ
Normative Expectation: Low Share	Inappropriate	6 (5.56%)	72 (66.67%)	78 (72.22%)
	Appropriate	4 (3.70%)	26 (24.07%)	30 (27.78%)
	Σ	10 (9.26%)	98 (90.74%)	108

7.4 Future Research & Limitations

Our study carries three main potentials for future research. First, concerning empirical expectations, participants in this study were only asked whether they expect most others to provide a fair share (dummy variable: yes/no). However, assessing the expected distribution would have been interesting for a more detailed analysis (e.g., "how many other participants, in percentage terms, do you think decided in favour of a fair split?"). The advantage of asking for a concrete distribution would have been to gain more detailed information about the distribution and how variations in the elicitation procedure of normative expectations may affect empirical expectations.

A second potential arises from assessing social norms in a specific cultural context. Social norms and the responses regarding the variation of the norm elicitation procedure may vary between different cultures. This study was conducted in the UK, and while its results concerning the effects of the elicitation method may extend/transfer to other Western cultures, repetition could test these finding in other cultures and regions.

Third, our results only hold for behaviours with a predominantly unambiguous social norm – in our context, fairness. Future research might investigate whether results also apply to behaviours with a more unambiguous or even misperceived social norm. Other social norms might not be as salient as the fairness norm is for the budget division.

7.5 Conclusion

The two-step norm elicitation method pioneered by Bicchieri and Xiao (2009) provides an easy-to-implement technique that captures personal normative beliefs and normative expectations about behaviours. Asking about the appropriateness of behaviour on a four-point scale provides researchers with more nuanced results, especially for detecting pluralistic ignorance and the misperception of beliefs. This study investigated

the sensitivity of variations in eliciting fairness beliefs and their impact on behaviours in a dictator game, using the two-step norm elicitation method. Running an online experiment with 438 subjects, we tested three different treatments. In line with previous studies, our results support the idea that incentives do make a difference. More specifically, incentivizing normative expectations significantly increases their correct estimation of the majority norm. By contrast to previous studies, the results of our experimental design indicate robustness for the time of elicitation: eliciting beliefs before the decision task did not influence behaviour compared to elicitation after the decision. The explanation might be two-fold: first, the rate of participants providing a fair share is already high across all treatments (on average, 66.67%). Second, the social norm of fairness is potentially well-known. Hence, in this study, elicitation before the task may not have had a priming effect, as the social norm already has been salient. This might be different in a more complex context, or when beliefs and behaviours are elicited about norms that entail more uncertainty. Finally, asking participants about the appropriateness of both a fair share and a low share did not significantly increase personal normative beliefs nor normative expectations. However, it did significantly increase empirical expectations about a fair share. Finally, the elicitation of normative expectations of both, a fair share and a low share, revealed that the fairness norm is very unambiguous, but one-quarter falsely believed that the majority regarded both shares as appropriate. The insights from this study may guide researchers who want to use the two-step norm elicitation method.

CHAPTER 8 | Conclusion

8.1 Overall Concluding Remarks

A set of studies in this dissertation reviewed some emerging concepts of how corporations can responsibly and sustainably digitalise. Another set of studies experimentally investigated social norms in internal whistleblowing applications and the robustness of a norm elicitation method. Even though research on Corporate Sustainability, Corporate Social Responsibility, whistleblowing, and social norms have proliferated for decades, the studies presented here identified important research gaps challenging how corporations and individuals contribute to sustainable development in a broader sense. Each chapter provides a conclusion and specific implications, respectively. In addition to that, a short and superordinate resume will be made, highlighting some key findings before pointing to future research avenues.

We learned from the literature reviews in Chapters 2 and 3 how Corporate Sustainability (CS) and Corporate Social Responsibility (CSR) are evolving in the context of digitalisation. Chapter 2 highlighted that digitalisation can be an enabler for CS. However, from a holistic sustainability perspective, it becomes crucial to enlarge the currently predominant environmental and economic efficiency-driven focus to integrate social and governance dimensions. Anchoring companies' sustainability and digital endeavours into the core strategy is needed to achieve a genuinely dual transformation (Epp et al., 2024). Moreover, Chapter 3 outlines the research on Corporate Digital Responsibility (CDR) and discusses whether the concept is warranted or redundant compared to studies on CSR intersecting digitalisation. Overall, we learned that CDR subsumes essential debates, for instance, on challenges of AI or data privacy and points to avenues like digital trust when signalling credible organisational responsibility in the digital era. CDR merges interdisciplinary debates of information systems and business ethics. Thereby, it strengthens the normative debate on responsibility in the realm of digitalisation more strongly than the current CSR literature, which focuses more predominately on measuring interrelations and performance. It is important to reintegrate CDR into CSR and CS approaches to pursue a unified strategy that aligns the responsibilities, challenges and potentials of digital technologies in terms of sustainable transformation.

After all, not walking the talk bears the risk of being perceived as machinewashing, described as a misleading use of AI and algorithm-based systems (Bernini et al., 2024), or ethical washing. Hence, institutionalising substantial CSR practices such as tangible corporate actions, policies, or initiatives is decisive. Providing potential internal whistleblowers with a secure mechanism to report misconduct is a means to institutionalise an ethical organisational culture and CSR commitment. Despite the necessity of increasing the institutional mechanisms of whistleblowing, we have seen in a review in Chapter 4 that coworkers can have a decisive influence on an individual's whistleblowing action and that whistleblowers experience adverse consequences from peers in the aftermath of whistleblowing.

As a result, the topic of social norms was addressed more intensively. To further narrow down some of the causal influences, we have learned from experimental studies in Chapters 5 and 6 about the critical role that the expectations about peers' perceptions have on whistleblowers. Findings demonstrate that whistleblowing behaviour is deeply influenced by normative expectations of peers—with misperceptions often deterring reporting and targeted social interventions showing promise in mitigating these effects. Chapter 5 uncovered misperceptions by revealing the divergence between personal normative beliefs and normative expectations. It further outlined the positive impact of social information interventions in correcting misperceptions. Chapter 6 more concretely investigated multiple expectations about all behaviour choices and found that the reporting behaviour is influenced not only by the normative expectations of reporting but also by the normative expectations of staying silent. However, revealing the actual majority norm that staying silent is regarded as inappropriate has a more substantial impact than revealing the majority norm regarding reporting. Finally, Chapter 7 contributes valuable methodological insights through an experiment investigating the robustness of the two-step norm elicitation procedure in experimental research. Findings provide researchers who want to apply the elicitation method with vital implications and valuable guidance.

To summarise, this dissertation contributes to theoretical, empirical, and methodological understanding across the fields of social norms and social responsibility using managerial and economic approaches. The insights gained underscore the importance of strategic alignment in corporate sustainability, the conceptual clarity in emerging concepts (CDR), the differentiated role of social norms in decision-making in

the context of whistleblowing, and the methodical robustness in measuring normative expectations.

8.2 Future Research Opportunities

The findings of the outlined studies pave the way for various future research. Five possible research paths and ideas will be emphasised without claiming completeness. Given the abundance and diversity of topics, each chapter outlines a comprehensive and specific presentation of the future research agenda, respectively. In addition, some overarching approaches are highlighted below. First, future research should continue to refine intersections, integrating multidisciplinary approaches. For instance, the intersection of corporate sustainability with digitalisation bears many ongoing research potentials. As outlined, there is a need to investigate further the strategic integration and multi-level approaches that extend beyond efficiency-driven perspectives. In this regard, it would also be worthwhile elaborating on how the dual transformation can be addressed and implemented, for instance, in terms of employee acceptance or identifying how social norms influence this change process, even though the combined approach, accounting for digital and sustainability dimensions, might be challenging. Similarly, in the context of CDR and CSR, the intersection with digitalisation requires multidisciplinary approaches combining digital technologies, technology ethics, corporate governance, as well as social, environmental and economic sustainability. Future research may explore how corporations can integrate these endeavours into a balanced ESG strategy. Stakeholder theory, particularly stakeholder engagement (Kujala et al., 2022), provides a promising approach to investigate how corporations can navigate digitalisation and CSR/CDR with stakeholders. It could further enable insights into how corporations account for the diverse expectations in a time of constant change driven by regulation and the speed of technological advances.

Second, whistleblowing research is an ongoing research area, and the potential influences that digital technology and artificial intelligence may have on whistleblowing are particularly interesting. Therefore, some exemplary questions will be raised subsequently. For instance, could corporations install automated fraud detection systems, and what would the ethical implications be? Do blockchain technologies provide a remedy in ensuring anonymity and protection, thereby making whistleblowers less vulnerable? How do multinational corporations balance local norms with global

whistleblowing standards? How does social proximity influence potential whistleblowers? Is there a difference between remote-working employees and those at the workplace regarding the likelihood of whistleblowing, thereby requiring a differentiated mechanism? Findings of research fields, such as human-machine interactions, may provide some vital possibilities to be incorporated when investigating digital whistleblowing.

Third, while social norms in this dissertation have been investigated in the application of whistleblowing, they may also provide explanatory power for other behaviours in the realm of CSR and sustainable development. There is, for instance, extensive literature on how social norms influence climate change risk perception (e.g., Van Der Linden, 2015), sustainable consumption behaviour (e.g., Pristl et al., 2021) or how CSR itself may become a social norm to which socially responsible human resource management can contribute (e.g., Shen & Benson, 2016). Future research may investigate whether and how social norms influence employees' adoption of CSR initiatives or the acceptance of transformational processes.

Fourth, social norms are not easily changed, but as illustrated in the whistleblowing application, potential misperceptions could be resolved by implementing social information interventions, thereby inducing behavioural change. Given their impact on behaviour, uncovering other areas where misperception may drive unethical behaviour, such as bribery and corruption or discrimination and toxic work culture, might be worthwhile.

A final suggestion for a research path would have been to provide a comparative overview of the different social norm elicitation techniques, their advantages, limitations, procedures and robustness checks. A recently published study shows just how timely and important this research is: Charness, Dimant, Gneezy and Krupka published an overview entitled 'Experimental Methods: Eliciting and Measuring Social Norms' on 29 March 2025 (still in press). Their study strengthens the motivation for the experiment conducted in Chapter 7 all the more, as they concluded that the method – referred to here as sequential opinion-matching – is "[n]ot yet systematically validated, frequently used both with [and] without monetary incentives" (Charness et al., 2025: 17). A research gap that the study in Chapter 7 closes. However, due to its topicality, the study by Charness et al. (2025) was not referenced in Chapter 7.

APPENDIX

Appendix Chapter 2 – From Digital Drift to Sustainable Direction: Reviewing the Interrelation of Corporate Sustainability and Digitalisation

Table A1: Overview of Reviewed Studies

Autor & Year	Title	Thematic Foci	Method (Data Collection)	Sustainability Dimension			Direction of Digitalization	
				Ecology	Economy	Social	Digital Sustainability	Sustainable Digitalization
Acciarini et al., 2021	Can digitalization favour the emergence of innovative and sustainable business models? A qualitative exploration in the automotive sector	innovative business models	case study	x	x	x	x	
Agrawal et al., 2023	Integration of artificial intelligence in sustainable manufacturing: current status and future opportunities	manufacturing; AI	SLR (n=196)	x	x		x	
Agrawal et al., 2023	Opportunities for disruptive digital technologies to ensure circularity in supply Chain: A critical review of drivers, barriers and challenges	circular economy; Industry 4.0	SLR (n=187)	x	x	x	x	
Al-Khatib, 2023	The impact of industrial Internet of things on sustainable performance: the indirect effect of supply chain visibility	Industry 4.0; manufacturing	questionnaire (n=380)	x	x	x	x	
Allal-Chérif et al., 2023	Born to be sustainable: How to combine strategic disruption, open innovation, and process digitization to create a sustainable business	sustainable entrepreneurship	case study	x	x	x	x	
Ardito et al., 2021	The duality of digital and environmental orientations in the context of SMEs: Implications for innovation performance	innovation	questionnaire (n=369)	x	(x)		x	x
Ardito, 2023	The influence of firm digitalization on sustainable innovation performance and the moderating role of corporate sustainability practices: An empirical investigation	innovation	survey (n=14,125)	x	x	x	x	
Bag et al., 2021	Industry 4.0 and supply chain sustainability: framework and future research directions	supply chain; Industry 4.0	two SLRs (n=242, n=76)	x	x	x	x	x
Belhadi et al., 2021	Analyzing the mediating role of organizational ambidexterity and digital business transformation on industry 4.0 capabilities and sustainable supply chain performance	Industry 4.0; circular economy	questionnaire-based survey (n=306)	x	x	x	x	x
Benešová et al., 2021	Design of a business readiness model to realise a green industry 4.0 company	Industry 4.0; environmental sustainability; maturity	literature analysis; questionnaire survey (n=50)	x			x	x

Böttcher et al., 2023	Digital sustainable business models: Using digital technology to integrate ecological sustainability into the core of business models	business models	collected case sample (n=31); literature analysis	x	x		x	
Broccardo et al., 2023	The interlink between digitalization, sustainability, and performance: An Italian context	performance; KPI	questionnaire (n=116)	x	x	x	x	
Chatzistamoulou, 2023	Is digital transformation the Deus ex Machina towards sustainability transition of the European SMEs?	twin transition	cross-sectional survey (n=23,464)	x	x	x	x	
Ching et al., 2022	Industry 4.0 applications for sustainable manufacturing: A systematic literature review and a roadmap to sustainable development	manufacturing; Industry 4.0	SLR (n=32)	x	x	x	x	
Contini et al., 2023	Developing key performance indicators for monitoring sustainability in the ceramic industry: The role of digitalization and industry 4.0 technologies	manufacturing; Industry 4.0; KPI	case study	x	x	x	x	
Cwiklicki & Wojnarowska, 2020	Circular Economy and Industry 4.0: One-Way or Two-Way Relationships?	circular economy; Industry 4.0	SLR (n=32)	x	x		x	x
Demir et al., 2023	Readiness and Maturity of Smart and Sustainable Supply Chains: A Model Proposal	supply chain	literature analysis; case study	x	x	x	x	x
Di Maria et al., 2020	Industry 4.0 technologies and circular economy: The mediating role of supply chain integration	Industry 4.0; circular economy; supply chain; manufacturing	questionnaire and database survey (n=1,229)	x	x		x	
Dwivedi & Paul, 2021	A framework for digital supply chains in the era of circular economy: Implications on environmental sustainability	circular economy	literature analysis	x			x	x
Dwivedi et al., 2023	Antecedents of digital supply chains for a circular economy: a sustainability perspective	circular economy; digital supply chain	literature analysis; experts' opinion (n=11)	x	x	x	x	
Feroz et al., 2023	Identifying organizations' dynamic capabilities for sustainable digital transformation: A mixed methods study	environmentally sustainable digital transformation	SLR literature review (n=195); questionnaire survey (n=63)	x			x	
Ferreira et al., 2023	Industry 4.0 implementation: Environmental and social sustainability in manufacturing multinational enterprises	social and environmental sustainability; Industry 4.0	interview survey (n=764)	x		x	x	
Findik et al., 2023	Industry 4.0 as an enabler of circular economy practices: Evidence from European SMEs	circular economy; Industry 4.0	interview survey (n=15.404)	x	x		x	
George et al., 2021	Digital Sustainability and Entrepreneurship: How Digital Innovations Are Helping Tackle Climate Change and Sustainable Development	ecosystems; innovation and design	interview	x	x	x	x	

Ghobakhloo et al., 2021	Industry 4.0, innovation, and sustainable development: A systematic review and a roadmap to sustainable innovation	sustainable innovation; manufacturing	literature review (n=70)	x	x	x	x	
Ghobakhloo et al., 2023	Actions and approaches for enabling Industry 5.0-driven sustainable industrial transformation: A strategy roadmap	Industry 5.0	content-centric literature review (n=91)	x	x	x	x	
Ghobakhloo et al., 2023	Intelligent automation implementation and corporate sustainability performance: The enabling role of corporate social responsibility strategy	intelligent automation; corporate social responsibility strategy	questionnaire survey (n=207)	x	x	x	x	x
Goede, 2021	Sustainable business intelligence systems: Modelling for the future	business intelligence systems	literature analysis	x	x	x	x	x
Grunwald, 2022	Sustainability co-creation in digitalized global value chains	stakeholder integration	literature analysis	x	x	x	x	
Guandalini, 2022	Sustainability through digital transformation: A systematic literature review for research guidance	digital transformation; sustainability	SLR (n=153)	x	x	x	x	
Gupta & Singh, 2021	Applications of emerging technologies in logistics sector for achieving circular economy goals during COVID 19 pandemic: analysis of critical success factors	circular economy; Industry 4.0	SLR (n=88)	x	x	x	x	
Haftor & Climent, 2021	CO2 reduction through digital transformation in long-haul transportation: Institutional entrepreneurship to unlock product-service system innovation	transportation industry and logistics	longitudinal case study	x	x		x	x
He et al., 2023	Driving mechanism model of enterprise green strategy evolution under digital technology empowerment: A case study based on Zhejiang Enterprises	green strategy	multiple case method (n=11)	x	x		x	
Histrov & Appolloni, 2021	Stakeholders' engagement in the business strategy as a key driver to increase companies' performance: Evidence from managerial and stakeholders' practices	value creation; stakeholder	case study (n=61)	x	x	x	x	x
Isensee et al., 2020	The relationship between organizational culture, sustainability, and digitalization in SMEs: A systematic review	environmental sustainability; digitalisation	SLR (n=80)	x			x	x
Islam et al., 2022	Annexing a Smart Sustainable Business Growth Model for Small and Medium Enterprises (SMEs)	design and innovation	literature review (n=355)	x	x	x	x	x
Jamwal et al., 2022	Deep learning for manufacturing sustainability: Models, applications in Industry 4.0 and implications	sustainable manufacturing; deep learning	literature analysis	x	x		x	

Kristoffersen et al., 2020	The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies	circular economy; manufacturing	SLR (n=32)	x	x		x	
Kristoffersen et al., 2021	The effects of business analytics capability on circular economy implementation, resource orchestration capability, and firm performance	circular economy; business analytics	survey (n=125)	x	x		x	
Kumar et al., 2021	Big data analytics application for sustainable manufacturing operations: analysis of strategic factors	clean technologies; environmental policy	literature analysis	x	x		x	
Kumar et al., 2022	Factors Influencing the Implementation of Industry 4.0 for Sustainability in Manufacturing	sustainable manufacturing	literature analysis; questionnaire-based survey (n=146)	x	x	x	x	
Lerman et al., 2022	Smart green supply chain management: a configurational approach to enhance green performance through digital transformation	supply chain	survey (n=473)	x		x	x	
Li et al., 2022	Digital Technology Adoption and Sustainable Development Performance of Strategic Emerging Industries: The Mediating Role of Digital Technology Capability and the Moderating Role of Digital Strategy	innovation; digital and sustainable performance	survey (n=385)	x	x		x	
Li, 2022	Digital transformation and sustainable performance: The moderating role of market turbulence	Digital transformation; sustainable performance	survey (n=223)	x	x		x	
Liu et al., 2022	A framework of digital technologies for the circular economy: Digital functions and mechanisms	circular economy	SLR (n=174)	x	x		x	
Lopes de Sousa Jabbour et al., 2018	Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations	circular economy; Industry 4.0	literature analysis	x	x		x	
Mukhuty et al., 2022	Strategic sustainable development of Industry 4.0 through the lens of social responsibility: The role of human resource practices	Industry 4.0; HRM; social responsibility	ILR (n=192)	x	x	x		x
Nayal et al., 2021	Supply chain firm performance in circular economy and digital era to achieve sustainable development goals	circular economy	questionnaire-based survey (n=297)	x	x		x	
Neligan et al., 2023	Circular disruption: Digitalisation as a driver of circular economy business models	business models; manufacturing	survey (n=599)	x	x		x	
Neri et al., 2023	What digital-enabled dynamic capabilities support the circular economy? A multiple case study approach	circular economy	case study (n=11)	x	x	x	x	

Niehoff, 2021	Aligning digitalisation and sustainable development? Evidence from the analysis of worldviews in sustainability reports	Industry 4.0; sustainability worldviews	case study (n=20)	x	x	x	x	x
Okorie et al., 2023	Digital transformation and the circular economy: Creating a competitive advantage from the transition towards Net Zero Manufacturing	circular economy; net-zero manufacturing emissions	literature analysis; engaged scholarship	x	x		x	
Pan & Nishant, 2023	Artificial intelligence for digital sustainability: An insight into domain-specific research and future directions	digital sustainability	SLR (n=41)	x	x	x	x	
Parmentola et al., 2022	Is blockchain able to enhance environmental sustainability? A systematic review and research agenda from the perspective of Sustainable Development Goals (SDGs)	blockchain usability	SLR (n=195)	x			x	x
Patil et al., 2023	Big data-Industry 4.0 readiness factors for sustainable supply chain management: Towards circularity	supply chain; big data; Industry 4.0; circular economy	SLR (n=146)	x	x	x	x	
Pauliuk et al., 2022	Co-design of digital transformation and sustainable development strategies - What socio-metabolic and industrial ecology research can contribute	co-design of sustainability and digitalisation	literature review	x			x	
Pînzaru et al., 2022	Adopting Sustainability and Digital Transformation in Business in Romania: A Multifaceted Approach in the Context of the just Transition	fair transition	questionnaire survey (n=128)	x	x	x	x	
Rejeb & Rejeb, 2020	Blockchain and Supply Chain Sustainability	logistics; supply chain; blockchain	SLR (n=79)	x	x	x	x	
Ribeiro et al., 2021	DSI Strategy Canvas: Modelling the Digital Social Innovation Strategy	digital social innovation	literature analysis	x		x	x	
Rusch et al., 2021	Application of digital technologies for sustainable product management in a circular economy: A review	circular economy	SLR (n=186)	x	x	x	x	
Sætra, 2022	The AI ESG protocol: Evaluating and disclosing the environment, social, and governance implications of artificial intelligence capabilities, assets, and activities	AI ESG protocol	literature analysis	x	x	x	x	x
Sahu et al., 2021	Integrating Industry 4.0 and circular economy: a review	circular economy; Industry 4.0	SLR (n=204)	x	x		x	
Sahu et al., 2023	Laminating STRATH block chain technology-SWOT architectures to endure business strategy between digital transformation, firms and supply chains capabilities for sustainability	blockchain; supply chain	literature analysis; experts' opinion	x	x	x	x	

Santarius & Wagner, 2023	Digitalization and sustainability: A systematic literature analysis of ICT for Sustainability research	ICT; ecological sustainability	SLR (n=215)	x	x	x	x	
Tasleem et al., 2019	Impact of technology management on corporate sustainability performance: The mediating role of TQM	corporate sustainability performance; total quality management; technology management	questionnaire (n=209)	x	x	x	x	
Torrent-Sellens et al., 2022	Boosting environmental management: The mediating role of Industry 4.0 between environmental assets and economic and social firm performance	environmental asset management; Industry 4.0; economic and social performance	panel data survey (n=1028)	x	x	x	x	
Ukko et al., 2019	Sustainability strategy as a moderator in the relationship between digital business strategy and financial performance	sustainability strategy	questionnaire (n=280)	x	x	x	x	
Wang et al., 2023	Investigating the impact of digital orientation on economic and environmental performance based on a strategy-structure-performance framework	manufacturing; supply chain; economic and environmental performance	questionnaire (n=300)	x	x		x	
Wei et al., 2022	How eco-control systems enhance carbon performance via low-carbon supply chain collaboration? The moderating role of organizational unlearning	low-carbon supply chain collaboration	questionnaire (n=297)	x	x		x	
Xu et al., 2023	Impacts of digital transformation on eco-innovation and sustainable performance: Evidence from Chinese manufacturing companies	manufacturing; sustainable performance; eco-innovation	questionnaire (n=210)	x	x	x	x	x
Yadav et al., 2023	Achieving the sustainable development goals through net zero emissions: Innovation-driven strategies for transitioning from incremental to radical lean, green and digital technologies	net-zero emissions	literature analysis	x			x	x
Zarte et al., 2022	Knowledge framework for production planning and controlling considering sustainability aspects in smart factories	sustainable manufacturing; product life cycle	SLR (n=153)	x	x	x	x	
Zheng et al., 2023	Leveraging technology-driven applications to promote sustainability in the shipping industry: The impact of digitalization on corporate social responsibility	CSR	Unbalanced panel data survey (n=28)	x	x	x	x	
zu Knyphausen-Aufseß & Santarius, 2021	Strategic management, the theory of the firm, and digitalization: reintroducing a normative perspective	ecosystems	literature analysis	x	x	x	x	x

Appendix Chapter 3 – Navigating Responsibility in the Digital Age: Systematic Literature Review Comparing CDR and CSR

Table A1: Literature Reviews on CDR

Authors & Year	Research Objectives	Journal / Area	Approach / Main results	Relationship with CSR
Aldboush & Ferdous, 2023	What are the ethical and privacy issues in fintech and how can they be handled?	Finance	SLR, ethical and privacy issues related to fintech, including bias, discrimination, privacy, transparency, justice, ownership, and control	n/a
Atanasov et al., 2023	Identify the current state of research on the integration of digital technologies in CSR activities in business	Risk and financial manageme nt	The analysis reveals that the intersection between digital technologies and CSR is corporate digital responsibility (model of: CDR, CSR activities, digital technologies)	Overlapping concepts
Bednárová & Serpeninova 2023	Review insight into the trends in the current literature	Digital accounting	Summary of definitions; subject areas, further descriptives (bibliometric analysis)	CDR as an extension of CSR
Carl, 2023	Review current research in Information Systems (IS) regarding CDR	IS	Scoping review on IS research, need for a more comprehensive view on data privacy and security	Independent concepts
Carl & Hinz, 2024	What is the state of research on CDR in information systems research?	IS	10 dimensions, 2-3 corresponding sub-dimensions, and several fields of action for each	Both overlap but are independent concepts
Covucci et al., 2024	What is the connection/distinction between Digital Sustainability and CDR?	Technology in Society	Dual-track systematic literature review (bibliometric and content analyses); collecting and comparing definitions and key topics of Digital Sustainability (DS) and CDR; DS often overlooks ethical implications, while CDR neglects broader sustainability impacts	CDR as an extension of CSR
Knopf & Pick, 2023	What definitions and underlying concepts of CDR are given in the scientific literature? Can a unified definition be reached across contexts?	Innovation	SLR, evaluate existing approaches, definitions, entrepreneurial motivation for CDR activities and consequences for corporations, business relationships and society, research gaps	CDR as an extension of CSR
Mueller, 2022	What are the roots of and current debates on CDR?	BISE	Motivational background and conceptual roots of CDR, definitions, contributions, synthesizing two key domains of CDR – a content-oriented perspective on digital ethics and an instrumental perspective on governance	Distinct, independent concepts
Weber-Lewerenz & Traverso, 2024	How can CDR be measured in line with CSR?	Sustainable Develop-ment	Structured literature analysis on the measurement of CDR in relation to CSR in the context of the construction industry	Both overlap but are independent concepts
Yadav & Mishra, 2022	How to comprehensively understand various aspects of the CDR across various fields of knowledge in recent time frame?	(Sustainabl e) Manage-ment	CDR is being assumed to be a strong differentiating tool for organizations in order to strengthen the trust of stakeholders and either prevent or minimize the potential threats	n/a

Note: IS= Information Systems, BISE = Business and Information System Engineering

Table A2: Identified Articles About CDR

Author(s) & Year	Method (Sample)	Community	Research Questions	Relationship of CDR and CSR			Thematical category		
				Independent, distinct, incl. overlaps	Subarea, part of CSR	Extension, evolution of CSR	CDR Debate	AI Ethics	Data Privacy, Trust
Agafonova et al., 2021	Empirical quantitative	Management (Marketing); BE, IS	Do consumers trust companies that are socially oriented and transparent information? Does proactive marketing activities in the digital environment lead to the emergence of new social risks?			x			x
Aitken et al., 2021	Empirical quantitative	Management (Finance), BE	How can socially minded data-intensive innovation be pursued in the private sector?	x		x	x	x	x
Bernini et al., 2024	Mixed Method	BE	Is it possible to operationalise and measure machinewashing by analysing companies' reporting? How can machinewashing be measured in terms of intensity, impact on ethical/ sustainable issues and types of deceptions towards stakeholders?		x			x	x
Carl, 2021	Theoretical	BE	How to evaluate CDR activities at the company level, particularly focusing on privacy and data security?	x			x		x
Carl, 2022	Empirical qualitative	IS, BE	How can companies ethically communicate their data privacy and security practices in the context of evolving consumer expectations and responsibilities?			x			x
Carl et al., 2024	Mixed Method	IS, BE	How are consumer preferences regarding companies' concrete CDR activities and how can these preferences inform the operationalization of CDR in practice?	x					x
Carl et al., 2022	Theoretical	IS, BE	How can the applicability of the existing CSR standard to CDR to pave the way for CDR standardization in the future?	x			x		x
Cheng & Zhang, 2023	Empirical quantitative	BE	How can CDR be conceptualized, measured and implemented?	x	x	x	x		x
Clausen et al., 2023	Mixed Method	Management (HR), IS	Which digital wellbeing initiatives are offered by organizations and/or expected by (potential) employees? How might digital wellbeing initiatives influence organizational attractiveness?			x	x		
Dörr & Lautermann, 2024	Theoretical	BE	How can Societal CDR give CDR a broader societal perspective?			x	x		x
Elliott & Copilah-Al, 2024	Empirical qualitative	Management	What are obstacles that hinder CDR implementation and how can they be overcome by managers?	x			x	x	x

Elliott et al., 2021	Theoretical	BE, IS	How can digital societal harms be avoided in AI systems using CDR?	x				x	x
Famularo, 2023	Empirical qualitative	BE, IS, Management	How does CSR discourse on digital issues affect large ICT firms?	n/a					x
Girrbach, 2021	Theoretical	IS, Management	What is CDR in the context of blockchain usage in supply chain management?			x	x		
Herden et al., 2021	Mixed Method	Management	What are the perceived opportunities and threats associated with the topic of digitalization?			x	x	x	x
Hartley et al., 2024	Theoretical	BE	What are the costs and benefits of a firm's CDR culture?	n/a			x	x	x
Jelovac et al., 2022	Theoretical	IS	How does CDR influence digital trust?	x					x
Jones & Comfort, 2021	Theoretical	Other (Gambling)	What effect does CDR have on sports betting?	n/a					x
Kärpänen, 2022	Empirical qualitative	BE, IS	What is the current state of knowledge regarding accessibility requirements and legislation in the context of CDR among micro-entrepreneurs in Finland?	x			x		x
Kluiters et al., 2023	Empirical quantitative	BE	How can a firm measure DT? What are the effects of firm- and governance-specific characteristics on DT? What are the effects of DT on firm value?	x					x
Kolyperas et al., 2024	Empirical qualitative	BE	How does CSR evolve and develop in the dynamic digital industry of esports?			x	x		
Kunz & Wirtz, 2024	Theoretical	Management (Marketing)	What is CDR and how does AI advancement affect it?	n/a			x	x	x
Liyanaarachchi et al., 2021	Theoretical	Management (Bank)	How can consumer data vulnerability in online banking be minimized by market-oriented CDR?		x		x		x
Lobschat et al., 2021	Theoretical	BE	How can ethical concerns in digital technologies and related data be handled?	x			x	x	x
Londoño-Cardozo & De Paz, 2021	Empirical qualitative	BE, IS, Management	How to effectively integrate digital contexts into traditional corporate social responsibility frameworks?		x		x		
Merbecks, 2024	Empirical qualitative	BE	How do companies report on their CDR-initiatives in Germany?	x			x		
Mihale-Wilson et al., 2022	Theoretical	IS	Why do we need CDR in addition to CSR? Does not the established CSR concept cover CDR as well?	x			x	x	

Mogaji et al., 2023	Theoretical	IS, Management	Introduces the concept of immersive time with the aim of working toward quantifying and qualifying the level of engagement in the metaverse.	n/a					(x)
Nagano, 2023	Theoretical	BE	How can institutional values contribute to CDR in ICT companies and how can these values address WEEE issues at multiple levels?	x	x	x	x		
Napoli, 2023	Empirical quantitative	Management	What are the relationships between the composition of a board of directors, digital technologies, CDR and greenhouse gas emissions?	x			x		
Orbik & Zozul'aková, 2019	Theoretical	Management	What is the relationship between CSR and digital transformation?	x			x		
Paluch et al., 2024	Theoretical	Management	What is the state of research on CDR?	x			x	x	
Pappas et al., 2023	Empirical qualitative	IS	The article highlights the emergence of CDR and the shift from industry 4.0 to industry 5.0, which focuses on human-centric approaches and human-AI partnerships.			x		x	x
Paul et al., 2024	Theoretical	Management	What is the state of research on digital transformation?	n/a			x	x	x
Pelters, 2021	Theoretical	BE, IS, Management,	How can digitalization be integrated into an existing theoretical sustainability construct? Can the BMJV's scenario technique be transferred to the university context?			x	x		
Peshkova, 2022	Empirical qualitative	Management (HR), IS, BE	The article investigates the implications and consequences of using digital footprint analysis technology in the recruitment process.	n/a					x
Rugeviciute, 2023	Theoretical	Management, IS, BE	How can organizations make responsible strategic decisions regarding the socio-environmental impacts of their ICT practices under the framework of CDR?			x	x		
Scarpi & Pantano, 2024	Theoretical	Management (Retail)	How can CDR be applied in AI retail service?	n/a				x	
Schneider, 2022	Mixed Method	BE, IS	How can CDR fill the accountability gap in business-to-government data sharing?			x			x
Schrödter, & Weissenberger, 2024	Empirical quantitative	BE	How do companies incorporate digital compliance as part of CDR?		x				x
Sidaoui et al., 2024	Theoretical	Management (Communi-cation)	How are organizational sensemaking processes of creation, interpretation, and enactment triggered by conversational AI issues and events?	n/a				x	
Stahl, 2024	Theoretical	BE	How can digital ecosystems be rendered responsibly?			x	x	x	
Suchacka, 2019	Mixed Method	BE, IS	How to characterise new challenges in CDR and new research areas which emerge in that field for social sciences?			x		x	x

Suchacka, 2020	Theoretical	BE	How does the awareness of responsibilities relate to technological and digital development among entrepreneurs, scientists, and decision-makers?			x		x	
Tóth & Blut, 2024	Theoretical	Other (Finance, AI)	How can the AI accountability framework and CDR be implemented in financial services?	n/a				x	
Trier et al., 2023	Theoretical	IS, Management	Overview of existing contributions to attain DR in the IS discipline, and discussions on the role of responsibility at the individual, corporate and societal level.	n/a			x	x	
Trittin-Ulbrich, & Böckel, 2022	Empirical qualitative	Management	How do institutional entrepreneurs understand/use CDR for responsible digital innovation?	x			x		
Van Der Merwe & Al Achkar, 2022	Theoretical	Other (Public policy)	How can CSR and CDR mechanisms be used for implementing responsible data use?	x					x
Volchek et al., 2024	Mixed Method	BE, IS, Management	Develop recommendations for the CDR strategy and a mathematical model of coordination decisions regarding CSR.			x	x		
Volkov & Sidorenko, 2022	Theoretical	BE, Other (Law)	How does the lack of a unified scientific direction for corporate responsibility on digital platforms, necessitating collaborative efforts among scholars?		x		x		
Vo Thai et al., 2024	Empirical quantitative	Management	How do human capital and stakeholder engagement influence CDR strategies, and what influence do CDR strategies have on firm performance?	x			x		
Weber-Lewerenz, 2021	Empirical qualitative	Other (Construction engineering)	How shall an adequate ethical framework be designed to support digital innovations in order to make full use of the potentials of digitization and AI?	n/a				x	
Wirtz & Pitardi, 2023	Theoretical	Management (Marketing), IS, BE	How will intelligent automation, service robots, and AI reshape service products and their delivery, particularly focusing on the implications for service firms and their marketing strategies?	x				x	x
Wirtz et al., 2023	Theoretical	Management	What are the risks and their mitigations of CDR in service firms?	n/a			x	x	x
Wynn & Jones, 2023	Mixed Method	IS	How can CDR and its parameters be modeled in a simple way?		x		x	x	x

*Table A3: Identified Articles About CSR and Digital**

Autor & Year	Method	Community	Research Question	Thematical category		
				Inter-action CSR & Digital	Perfor-mance View	Digital Commu-nication
Abad-Segura et al., 2024	Empirical Quantitative	BE	How can integrating STI and CSR enhance competitiveness in the Spanish agricultural and livestock subsectors while addressing the inherent economic, social, and environmental challenges?	x	x	
Abdallah-Ou-Moussa et al., 2024	Mixed Method	Management, BE	What is the impact of digitalization on CSR in the automobile insurance sector in Morocco?	x		
Aleksić et al., 2024	Empirical Quantitative	BE	How do employees acknowledge meaningful work in circumstances that offer limited opportunities for meaning (e.g., in the context of the COVID-19 crisis)? How do situational events arise from digitalization and the COVID-19 context affect the three elements of the organizational frustration model?	x		
Alfalah et al., 2022	Empirical Quantitative	BE	How can the investment in IT and the corporate governance practices of a Saudi Arabian telecoms company impact the company's overall financial performance?		x	
Almeida et al., 2022	Empirical Quantitative	Management	What is the state of digital CSR in Portuguese companies of the water industry?		x	
Al-Omoush et al., 2024	Empirical Quantitative	BE	How did digital CSR impact social entrepreneurship, organizational resilience and competitive intelligence during COVID-19?		x	
Al-Omoush, 2024	Empirical Quantitative	BE	What role do institutional pressures and top management support play in digital CSR, and how does digital CSR impact social trust and corporate sustainability?		x	
Arnal-Pastor & Berné-Martínez, 2024	Empirical Qualitative	Other (Journalism)	How are social innovation and CSR portrayed in Spanish media?			x
Bhattacharyya, 2023	Empirical Qualitative	BE	How do online reviews of CSR initiative of firms reflected the logic used by managers in organizations?			x
Campoamor, 2019	Empirical Qualitative	BE	How can a Spanish telecom giant legitimizes its market expansion through a CSR narrative that links generic notions of technological innovation and children's rights to projects of development and democracy?			x
Chang et al., 2023	Empirical Quantitative	Management (Innovation)	How are different forms of innovation and CSR intertwined?	x		
L. Chen & Chen, 2023	Empirical Quantitative	Management (Innovation)	What is the relationship between Digital Innovation and CSR from the metaorganizational perspective?	x		
Chen, 2023	Empirical Quantitative	BE	What is the impact of China's digital economy development on low-carbon innovation and what is the value of CSR in the digital era?		x	
Djakman & Siregar, 2024	Empirical Quantitative	Management (Risk), BE	How is the effect of the maturity learn element in ERM and CSR as risk management on the level of digital transformation related? Does CSR moderate on the association between ERM and the level of digital transformation?	x		
Esposito & Ricci, 2021	Empirical Qualitative	BE	How are digitization processes influencing the attitude toward CSR in cultural organizations?			x

Etter et al., 2019	Theoretical	BE, IS, Management	What was once hailed as a sustainable and communal lifestyle movement rooted in counterculture, increasingly come to be seen as the posterchild for all that is wrong with contemporary capitalism	x		
Fu et al., 2023	Empirical Quantitative	Management (Consumer Studies), BE	Can green agricultural products consumption behaviour be enhanced through CSR information transparency on digital platforms and consumers' online identification?			x
Ghobakhloo et al., 2023	Empirical Quantitative	BE	How do firms' technological, organizational, environmental and HR contexts impact Intelligent Automation implementation, and how does it impact CSR performance?		x	
Gilbert et al., 2024	Theoretical	BE, IS	How can gatekeepers exhibit ethical responsibility in their efforts to gain, maintain, and sustain their moral legitimacy?	x		
González-Ramírez et al., 2024	Empirical Quantitative	BE	How are digitalisation, sustainability, business environments, innovation and CSR connected?	x		
Govindan, 2024	Mixed Method	BE	How do Industry 4.0 technologies integrate with CSR at the practitioner level to best achieve SDGs? What are the common practices involved in CSR 4.0 implementation in MSMEs under practice-based view? What is the most influential practice to effectively implement CSR 4.0 in companies under social good theory? What elements of the cause and effect analysis among common CSR 4.0 practices result in the most effective implementation?	x		
Guo et al., 2024	Theoretical	BE	This paper studies the governance of social responsibility of platform corporations from the perspective of social subjects in the context of platform transformation of traditional corporations.	x		
Hu & Liu, 2023	Empirical Quantitative	BE	How does digital technology development influence CSR in China?		x	
Huang & Shen, 2024	Empirical Quantitative	BE	Is there a bidirectional relationship between digital transformation and CSR? If so, does this relationship vary with different levels of economic policy uncertainty?	x		
Huang & Wei, 2023	Empirical Quantitative	BE	Does CEOs' green experience affect environmental CSR?	x		
Jiang et al., 2023	Empirical Quantitative	BE	Does digitization drive CSR?	x	x	
Jiménez et al., 2021	Theoretical	BE	CSR self-regulation instruments, codes of conduct			x
Jung et al., 2022	Empirical Quantitative	Management (Consumer studies)	How does perceived CSR and corporate social irresponsibility affect electronic "word of mouth" for large versus small companies?			x
Khan et al., 2023	Empirical Quantitative	(Sustainable) Management, BE	How do digital technologies, tax avoidance, and green employee behavior affect the sustainable performance of the firms? How does CSR moderate the relationships between DT, GEB, TA, and sustainable firm performance?		x	
Khattak & Yousaf, 2021	Empirical Quantitative	Management	How does digital social responsibility help achieve strategic performance and CSR performance?		x	x
Kong & Liu, 2023	Empirical Quantitative	BE	How does the adoption of digital technology shape CSR in China?	x		
Koutras, 2019	Theoretical	BE	How are commercial publishers' and authors' interest interconnected?			
B. Li et al., 2024	Empirical Quantitative	Other (Fintech)	How is Fintech business related to CSR?	x		

H. Li et al., 2024	Empirical Quantitative	BE	What is the relationship between digital innovation and CSR performance?		x	
M. Li, 2021	Empirical Quantitative	Management (Advertising)	What are synergistic effects of solutions journalism and CSR advertising?			x
Liao et al., 2023	Empirical Quantitative	(Sustainable) Management, BE	What is the relationship between ECSR/ICSR and BMI? Examination of value co-creation (VCC) as the mediating factor within the interconnection connecting CSR fulfillment and BMI, and the role of DPC in moderating these mechanisms' analyses.		x	
Lin et al., 2024	Empirical Quantitative	BE	How do digital orientation and CSR activities affect alliance relationship stability? How does relationship quality mediate these relationships?	x		
Lindman et al., 2023	Theoretical	BE	What types of power do large Internet tech companies have?	x		
H. Liu et al., 2024	Empirical Quantitative	BE	How has the digital economy era enhanced CSR in China?		x	
W. Liu et al., 2024	Empirical Quantitative	BE	Does the release of digital supply chain announcements disclosing CSR information by Chinese A-share listed firms lead to positive market reactions? Do CSR strategy type, CSR value type and CSR stakeholder orientation reflected in the announcements influence the market reaction to the firm's stock market? Do the embeddedness of supply chain relationship, supply chain digital breadth and digital depth reflected in the announcement affect the market reaction to the firm's stock market?		x	
López-Nicolás et al., 2024	Empirical Qualitative	Management (Innovation)	What is BM (Innovation) focus of FFs? What is the purpose of BMI in FFs? What combination/ configuration of components in FFs can lead to specific types of BMI? How does family governance play a role in the BMI process?		x	
Ma et al., 2021	Empirical Quantitative	BE	How does CSR engagement on social media (CSRS) influence electronic word-of-mouth (eWOM) among consumers? What role does consumer-company identification (CCI) play in mediating the relationship between CSRS and eWOM?			x
Y. Ma et al., 2024	Theoretical	BE	How do e-commerce platforms balance price discrimination with their CSR?	x		
McBride et al., 2019	Empirical Qualitative	BE, Other (Law)	What are the viable policy initiatives to regulate digital retouching in advertising? How can CSR be leveraged to reduce the harmful effects of digitally altered images? What legal frameworks exist in the US to support these initiatives?			x
Nie et al., 2024	Empirical Quantitative	BE, Management	Can engaging in CSR improve a firm's resource base (e.g., human, financial, and technical resources), thus promoting D-TRANF?		x	
Okazaki et al., 2020	Empirical Quantitative	Management (Marketing)	To what extent do brands that lead their industries in CSR programs use social media to broadcast their CSR efforts and establish relational and participative environments with consumers?			x
Özturan & Grinstein, 2022	Empirical Quantitative	BE, Management (Marketing)	Examining trends in tandem with a focus on global brand CMOs; compares the impact of CSR communication and sociopolitical activism communication in light of two factors: global brands' origin and CMOs' nationality			x
Pan et al., 2021	Empirical Quantitative	BE	How does customer orientation impact technology orientation on firms' NPD performance? How may firms' engagement in CSR moderate the relationships between different strategic orientations on NPD performance?	x	x	
Park et al., 2021	Mixed Method	BE, Management	What dimensions of CSR are frequently observed among successful CSR campaigns? Which CSR communication strategies are frequently observed among successful CSR campaigns? Which digital engagement strategies are frequently employed by companies to engage their stakeholders in the digital			x

		(Stakeholder relationships)	environment? How do companies use third-party endorsement and engage nonprofits, opinion leaders, and influencers in their CSR communication? What are the primary intended outcomes of CSR programs? What types of research are conducted to understand the stakeholder and situation? Which stakeholders are targeted for and engaged with CSR communication? What are the forms of traditional and digital media frequently utilized for CSR communications?			
Peng et al., 2024	Empirical Quantitative	Management (Communication)	What are the differences between enterprise- and co-generated content on consumer attitudes after negative Social Responsibility events?		x	x
Prisco et al., 2024	Empirical Quantitative	Management (Consumer studies)	Assess consumer attitudes and intentions regarding food delivery app usage			x
Purnamawati et al., 2023	Empirical quantitative	Other (Economic Development)	What is the role of the green economy and digitalization for sustainable village economic development with CSR as a moderating variable?	x		
Rangel-Pérez et al., 2023	Empirical Quantitative	(Sustainable) Management, BE	Is there a relationship between continuous investment in CSR and the share price of Spanish banks analysed by the MERCO ESG Responsibility Ranking, between 2011 and 2019 (without considering the effects of COVID-19)?		x	
Shen et al., 2024	Mixed Method	BE	How do CPV and CPCSR affect CCB in different shopping contexts? What are the differences in consumer behavior between online and offline shopping environments?			x
Shestakova, 2024	Theoretical	BE, Management (HR)	Analysing the historical evolution of the discussion on the place and role of man in the labour process in connection with the turns of scientific and technological progress and identify the distinctive features of the contemporary situation; Pointing out the new paradigmatic changes; Emphasising the need to strengthen the factor of social responsibility; Explaining some of the possible directions of socio-economic regulation	x		
Shkalenko & Nazarenko, 2024	Mixed Method	(Sustainable) Management, IS, BE	How can AI and IoT be integrated into CSR strategies to enhance financial risk management? What are the potential impacts of these technologies on sustainable development? How do institutional structures need to adapt to support the integration of AI and IoT into CSR practices?	x		
Srivetbodee & Igel, 2021	Empirical Qualitative	(Sustainable) Management, IS, BE	Examine IoT, AI, and big data technology as a means for improving agricultural productivity; identify the success factors and obstacles that corporations and farmers encounter when adopting smart technologies, and explore the impact of smart farming technologies on CSR performance in this sector.		x	
Stock et al., 2022	Empirical Quantitative	Management, BE	What is the relationship between CSR and an SME's digital innovation, thus presenting far-reaching implications for SME research and the emerging scholarly debate on digital innovation in resource-constrained organisations?	x		
Sun et al., 2024	Empirical Quantitative	BE, Management	Can digital transformation improve CSR performance? If it can, what is the impact mechanism? Does the impact of digital transformation on CSR performance vary in different scenarios?		x	
Thuong, 2024	Empirical Quantitative	BE, Management	How does the relationship between CSR and digitalization affect banking performance in Vietnam? How do ownership characteristics and listing status affect this relationship? Does the COVID-19 pandemic change the way CSR and digitalization affect banking performance?		x	
Tuyen et al., 2023	Empirical Quantitative	Management (Innovation)	Does corporate engagement in social responsibility affect firm innovation? The mediating role of digital transformation		x	
Vítová, 2022	Theoretical	BE	What is the connection between the unfair commercial practices and the concept of CSR from the perspective of consumer law?			x

Wang et al., 2024	Empirical Quantitative	BE	How is the effect of the occupational pension on CSR and organizational resilience?		x	
L. Wang & Yan, 2023	Empirical Quantitative	Management (Innovation), BE	How is the effect of DT and the firm's innovation performance and the boundary condition of CRS?		x	
M. Wang et al., 2024	Empirical Quantitative	BE	How is the intersection of CSR and public health within the context of digital platforms? How is the impact of digital platforms on the sustainable development practices of enterprises, seeking to comprehend how these platforms influence the implementation of environmental protection policies, resource management, and social responsibility initiatives?	x		
Wu et al., 2024	Empirical Quantitative	BE	What role does CSR in tech innovation play in enhancing the sustainable competitive performance of firms?		x	
Xin et al., 2022	Empirical Quantitative	BE	What is the impact of digital finance on CSR performance of pollution-intensive industry?	x		
Xu et al., 2023	Empirical Quantitative	BE	How does the digital transformation affect CSR in the mining industry?	x		
Yang & Jin, 2024	Empirical Quantitative	Management, IS, BE	This study enriches and expands the existing knowledge system on this topic by integrating theories related to the digital economy and resource-based theories.		x	
S. Yang et al., 2023	Empirical Quantitative	BE	How to practice social responsibility to make customers believe in the continuous value co-creation between the bank and its customers is an essential issue? What factors transmit the belief in this win-win situation?			x
Zheng et al., 2023	Empirical Quantitative	BE, Management	How do digitalization initiatives across the value chain impact a firm's CSR performance in the shipping industry?		x	
Y. Zheng & Zhang, 2023	Empirical Quantitative	Management, BE	What effects does the digital transformation have on green technology innovation within businesses? Activate or block? How do the efforts of green technology innovation impact enterprise digital transformation? How does a company's innovation in green technology depend on the extent of its digital transformation? Does the relationship between digital transformation and the innovation of green technologies depend on distinct geographical areas or property rights?		x	
Zhong & Ren, 2024	Empirical Quantitative	Management (Innovation)	How does digitalization affect firms' innovation efficiency, and does CSR matter?	x	x	
Zhou et al., 2024	Empirical Quantitative	BE, IS	How does structured data presentation influence users' engagements on WeChat and Facebook platforms in Chinese and Latin American firms? How can system quality be enhanced to promote the use of CSR programs, since it is still unclear how it mediates user interaction within the CSR communication framework? How can CSR programs improve the effects of well-synchronized data on user engagement, although there is an acknowledgment of the impact of CSR on various business performances, but the role of CSR regarding the relationship between data organization and user involvement is still unestablished?	x		x
Zhu et al., 2024	Empirical Quantitative	(Sustainable) Management, IS, BE	How do CEOs' digital technology backgrounds influence enterprise digital transformation? What is the mediating effect of R&D investment and CSR on the relationship between CEOs' DTBs and EDT? In what contexts do CEOs' DTBs have a more significant impact on digital transformation?	x		

Table A4: Definitions of Corporate Digital Responsibility

Author	Definition
Aitken et al., 2021	A means of centring digital responsibility within private sector practice, a set of norms and values to guide organisations' approaches to both creating and using digital technologies.
Carl, 2021	CDR seeks to ensure an ethical and responsible development, deployment, and use of digital technologies and data. CDR puts, inter alia, privacy and data security attempts in a broader context to provide a more holistic approach.
Carl et al., 2024	CDR comprises a set of principles designed to encourage the ethical and conscientious development, adoption, and utilization of digital technologies.
Dörr & Lautermann, 2024	Societal CDR is defined as the responsibility of companies to develop their digital business strategies considering the impacts on societal stakeholders and institutions.
Elliott et al., 2021	CDR is a voluntary commitment by organisations fulfilling the corporate rationalisers' role in representing community interests to inform 'good' digital corporate actions and digital sustainability (i.e. data and algorithms) via collaborative guidance on addressing social, economic, and ecological impacts on digital society.
Elliott & Copilah-Ali, 2024	In addition to Lobschaft et al. 2022: CDR is first and foremost a voluntary commitment by organisations in representing societal interests and to inform 'good' digital corporate actions and digital sustainability (i.e., using data and algorithms) via collaborative guidance on addressing social, economic, and ecological impacts on digital society.
Girrbach, 2021	assuming responsibility for economic, social, and ecological as well as digital aspects focusing on the chances for sustainable issues arising out of digitalization
Hartley et al., 2024	CDR is defined as the principles that guide a firm 's ethical, fair, and protective use of data and technology within their digital ecosystem
Herden et al., 2021	Corporate Digital Responsibility is an extension of a firm's responsibilities which takes into account the ethical opportunities and challenges of digitalization.
Jelovac et al., 2022	An assemblage of practices and behaviours that help an organization use data, digital technologies in a way that is socially, economically, technologically and ecologically responsible.
Lobschat et al., 2021	CDR is defined as a set of practices and behaviours that help an organisation use data and digital technologies in ways that are perceived as socially, economically, and environmentally responsible.
Mihale-Wilson et al., 2022	obligations that companies have towards society, focused on responsibility in relation to developing and using technology
Trittin-Ulbrich & Böckel, 2022	CDR emphasizes the voluntary, self-regulatory character of corporate commitment to responsible digital innovation.
Van Der Merwe & Al Achkar, 2022	We propose to define CDR as the set practices, policies, and governance structures of corporations as they relate to the digital transformation.
Weber-Lewerenz & Traverso, 2024	This study defines CDR as a corporate concept to meet CSR when dealing with digital technologies and AI in order to achieve a highly positive impact on people, society and the environment and to fulfill ESGs.
Wirtz et al., 2023	We define CDR in the context of service as the principles underpinning a service firm's ethical, fair, and protective use of data and technology when engaging with customers within their digital service ecosystem.

Table A5: Glossary Used to Identify the Digital Technologies Discussed in the Papers

Category	Definitions, Synonyms, and Abbreviations
Digital Technology Application	Digital marketing, online marketing, internet marketing, mobile internet, mobile web, industrial internet, Industry 4.0, IIoT (Industrial Internet of Things), e-commerce, online shopping, electronic commerce, mobile payment, M-payment, third-party payment, NFC (Near Field Communication) payment, contactless payment, B2B (Business-to-Business), B2C (Business-to-Consumer), C2B (Consumer-to-Business), C2C (Consumer-to-Consumer), O2O (Online-to-Offline), IoT (Internet of Things), smart agriculture, precision farming, AgriTech, smart investment, automated trading, robo-advisors, smart transportation, intelligent transport systems (ITS), autonomous vehicles, smart mobility, ride-sharing, MaaS (Mobility as a Service), smart service, digital services, cloud-based services, SaaS (Software as a Service), smart healthcare, digital health, telemedicine, e-health, mHealth, internet healthcare, telehealth, smart wearables, wearable tech, smart home, home automation, IoT home devices, smart environmental protection, environmental monitoring, green tech, smart cultural tourism, digital tourism, smart grid, intelligent grid, smart energy, green energy, renewable energy tech, smart marketing, AI marketing, data-driven marketing, fintech, financial tech, digital banking, mobile banking, neobanks, insurtech, regtech, digital assets, unmanned retail, cashier-less stores, automated checkout, digital finance, online finance, financial technology, internet finance, web-based finance, quantitative finance, algorithmic trading, high-frequency trading (HFT)
Artificial Intelligence (AI)	AI, artificial intelligence, cognitive computing, business intelligence (BI), data intelligence, intelligent data analysis, advanced analytics, image understanding, computer vision, intelligent robot, robotics, automation, autonomous systems, ML (machine learning), AI-driven decision-making, predictive analytics, investment decision support system, decision support systems (DSS), biometric technology, biometrics, fingerprint recognition, facial recognition, iris recognition, voice recognition, deep learning (DL), neural networks, speech recognition, voice AI, virtual assistant, semantic search, NLP (natural language processing), text analytics, automatic driving, self-driving cars, autonomous vehicles (AV), face recognition, facial biometrics, identity verification, authentication technology
Big Data	Big data, data science, data analytics, data engineering, data-driven decision-making, data lakes, data warehouses, data mining, knowledge discovery in databases (KDD), data visualization, dashboards, business intelligence tools, text mining, text analytics, NLP (natural language processing), heterogeneous data, unstructured data, semi-structured data, massive data, large-scale data processing, high-dimensional data, credit reporting, credit scoring, augmented reality (AR), virtual reality (VR), XR (Extended Reality), mixed reality (MR), 3D simulation, immersive technology
Cloud Computing	Cloud computing, cloud-based solutions, cloud services, SaaS (Software as a Service), PaaS (Platform as a Service), IaaS (Infrastructure as a Service), memory computing, in-memory computing, RAM-based computing, graph computing, knowledge graphs, stream computing, real-time analytics, edge computing, fog computing, multiparty secure computing, privacy-preserving computing, secure multi-party computation (SMPC), homomorphic encryption, green computing, sustainable computing, energy-efficient computing, brain-inspired computing, neuromorphic computing, cognitive computing, billion-level concurrency, high-performance computing (HPC), converged architecture, hyperconverged infrastructure (HCI), Internet of Things (IoT), industrial IoT (IIoT), cyber-physical systems (CPS), edge AI, smart sensors
Blockchain	Blockchain, distributed ledger technology (DLT), digital currency, cryptocurrency, virtual currency, tokenized assets, stablecoins, smart contract, self-executing contracts, decentralized applications (DApps), distributed computing, decentralized computing, decentralization, Bitcoin (BTC), Ethereum (ETH), altcoins, alliance chain, consortium blockchain, private blockchain, public blockchain, differential privacy technology, privacy-preserving blockchain, zero-knowledge proofs (ZKP), homomorphic encryption, consensus mechanism, proof-of-work (PoW), proof-of-stake (PoS), delegated proof-of-stake (DPoS), Byzantine fault tolerance (BFT), sharding, interoperability, cross-chain technology

Note: Based on Li, Lu, Lin & Meng (2024). Digital innovation and corporate social responsibility performance: Evidence from firms' digital patents. Technological Forecasting and Social Change, 207, 123626. Extended by OpenAI's ChatGPT-4-turbo (March 2025 version). Prompt to advance the original list: „You are a professional data scientist and an expert in digital technologies. To analyse which digital technologies are mentioned in a collection of 100 PDF files you aim to establish a dictionary. Use the following dictionary as a basis and add all synonyms and abbreviations that help to detect which digital technologies are mentioned in the PDF files. Establish a table that is much more detailed than the following: list by Li, Lu,Lin & Meng (2024).“

Appendix Chapter 5 – I don't believe that you believe what I believe: Experiment on Misperceptions of Social Norms and Whistleblowing

Appendix A: Survey

In this online appendix, we provide a more detailed analysis of the survey results. We start the evaluation of the survey by looking at the different types of wrongdoing observed by the participants (see Table A1). In the survey, 17% of the participants reported having observed abusive or intimidating behaviour towards other employees, 8% having observed conflicts of interest (e.g., decisions made or actions taken that benefit the respective employee, their family or friends, over the interests of the organization), 16% having observed peers that had been dishonest with other employees, customers, vendors, or the public, 2% observed someone offering bribes and/or inappropriate gifts, 13% observed violations of health and/or safety regulations, 9% observed retaliation against someone who reported misconduct and 27% observed stealing or theft. Six percent of the participants observed other kinds of wrongdoing and 2% did not specify the wrongdoing. For almost all types of observed wrongdoing, the majority holds the personal normative belief in support of whistleblowing, except for offering bribes and/or inappropriate gifts, and violations of health and/or safety regulations, for which the majority is personally in support of remaining silent. Regarding conflicts of interest, 50% are in support of remaining silent, and 50% are in support of whistleblowing. Interestingly, for all types of observed wrongdoing, survey participants assumed that the majority would support remaining silent (with the exception of offering bribes and/or inappropriate gifts, where the expected support is 50/50).

Table A1: Specification of the Type of Wrongdoing Observed by Participants

Type of wrongdoing	
Abusive or intimidating behaviour towards other employees	17%
Conflicts of interest	8%
Lying to employees, customers, vendors or the public	16%
Offering bribes and/or inappropriate gifts	2%
Violations of health and/or safety regulations	13%
Retaliation against someone who reported misconduct	9%
Stealing or theft	27%
Other kinds of wrongdoing	6%
Did not specify the wrongdoing	2%
Consequence for the organization	
Wrongdoing not disadvantageous for organization	36%
Wrongdoing disadvantageous for organization	53%
Did not know	11%

Thirty-six percent of survey participants indicated that the observed wrongdoing was neither disadvantageous to the organization nor did it harm the organization, 53% indicated that it was disadvantageous to the organization or harmed it, and 11% indicated that they did not know. If we compare only the participants who perceived that the wrongdoing was not disadvantageous to the organization with the participants who indicated that it was disadvantageous, we find that the latter were significantly more likely to report the wrongdoing ($\chi^2 = 4.2343$; $p = 0.040$; 22.22% vs. 43.40% report wrongdoing). We reveal a similar tendency for personal normative beliefs: Participants who perceived that wrongdoing was disadvantageous to the organization were significantly more likely to be in support of whistleblowing than participants who indicated that wrongdoing was not disadvantageous to the organization ($\chi^2 = 4.3262$; $p = 0.038$; 71.7% vs. 50% have a personal normative belief in support of whistleblowing). For normative expectations, we observe that 25% of the participants who indicated that the wrongdoing was not disadvantageous to the organization hold normative expectations in support of whistleblowing compared to 47.17% of participants who indicated that the wrongdoing was disadvantageous to the organization, which results in a significant difference ($\chi^2 = 4.4632$; $p = 0.035$).

In the next step, we focus on whether we observe a discrepancy between personal normative beliefs and normative expectations. For 66% of the subjects, there is no discrepancy between their personal normative beliefs (PNB) and their normative expectations (NE) regarding whistleblowing (see Figure A1). Twenty-six percent of the subjects personally believed that not reporting the wrongdoing was the correct course of action and expected the majority of their colleagues to share this belief. Forty percent of the subjects indicated that, personally, whistleblowing was for them the right action to do, and expected the same from their colleagues. However, for 34% of the subjects, we observe a discrepancy between personal normative beliefs and normative expectations regarding whistleblowing. For most of these, the discrepancy goes in the presumed direction: 25% of all participants held the personal normative belief that whistleblowing was the right thing to do when observing wrongdoing in their organisation, but they expected that the majority of their colleagues would believe that not reporting the wrongdoing was the right action. Only 9% of subjects who personally disapproved of whistleblowing anticipated that the majority of their colleagues would believe that it was the appropriate action when faced with observing wrongdoing.

Figure A1: Distribution of Discrepancy and Non-Discrepancy

	Personal normative belief: whistleblowing	Personal normative belief: silence
Normative expectations: whistleblowing	40%	9%
Normative expectations: silence	25%	26%
No Discrepancy between personal normative belief and normative expectations: **66%**		
Discrepancy between personal normative belief and normative expectations: **34%**		

Finally, we consider the relationship between personal normative beliefs, normative expectations, and whistleblowing behaviour (see Table A2). Both personal normative beliefs and normative expectations seem to have a significantly positive relationship with self-reported whistleblowing behaviour. In particular, among the 35 participants who held personal normative beliefs against whistleblowing, only 3 did blow the whistle, whereas among the 65 participants who held personal normative beliefs in support of whistleblowing, 33 reported the wrongdoing of their colleague. This difference is statistically significant ($\chi^2 = 17.582$; $p < 0.001$). Similarly, among the 51 subjects who expected their colleagues to be against reporting the wrongdoing as the right course of action, only 11 blew the whistle. Comparing this to the 49 participants who believed that their colleagues were in support of whistleblowing, of whom 25 actually became whistleblowers, reveals a significant difference ($\chi^2 = 9.408$; $p = 0.001$). This relationship between normative expectations and whistleblowing behaviour remains consistent if considering only participants who believe whistleblowing to be the right thing to do: Among the 25 participants who personally were in support of whistleblowing, but expected others to be against it, 8 actually blew the whistle. In contrast, among the 40 participants who both personally supported whistleblowing and expected their colleagues to do the same, 25 reported the wrongdoing, resulting in a statistically significant difference ($\chi^2 = 5.726$; $p = 0.009$). Table A3 shows the relationship between personal normative beliefs, normative expectations, and whistleblowing behaviour among female and male participants. Overall, there is no significant difference between gender whether in whistleblowing behaviour ($\chi^2 = 1.8141$; $p = 0.404$), in normative expectations ($\chi^2 = 2.2930$; $p = 0.318$), or in personal normative beliefs ($\chi^2 = 1.8166$; $p = 0.403$).

Table A2: Relationship between personal normative beliefs, normative expectations, and whistleblowing behaviour

		Behaviour		
		Silence	Whistleblowing	Σ
Personal normative belief	Silence	32	3	35
	Whistleblowing	32	33	65
Normative expectations	Silence	40	11	51
	Whistleblowing	24	25	49
Normative expectations (only participants who hold personal normative belief in support of whistleblowing)	Silence	17	8	25
	Whistleblowing	15	25	40

Table A3: Relationship between Personal Normative Beliefs, Normative Expectations, and Whistleblowing Behaviour, separated by Female and Male Participants

Female		Behaviour		
		Silence	Whistleblowing	Σ
Personal normative belief	Silence	13	2	15
	Whistleblowing	19	16	35
Normative expectations	Silence	18	5	23
	Whistleblowing	14	13	27
Normative expectations (only participants who hold personal normative belief in support of whistleblowing)	Silence	8	3	11
	Whistleblowing	11	13	24

Male		Behaviour		
		Silence	Whistleblowing	Σ
Personal normative belief	Silence	19	1	20
	Whistleblowing	13	16	29
Normative expectations	Silence	22	6	28
	Whistleblowing	10	11	21
Normative expectations (only participants who hold personal normative belief in support of whistleblowing)	Silence	9	5	14
	Whistleblowing	4	11	15

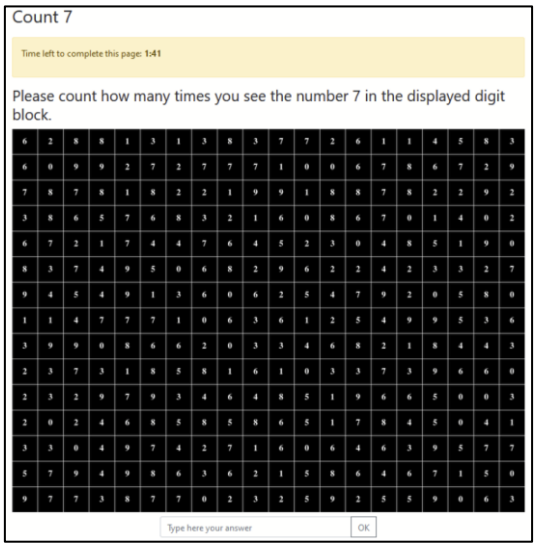
Appendix B: Instructions

General notifications

- In this experiment there are no other tools allowed except for the computer mouse and the keyboard.
- During the experiment all amounts are quoted in the **fictitious currency „token“.** All payouts achieved in this experiment will be exchanged at the end at an exchange rate of **0.50 euros per 1 token**. In addition, you will receive a show-up fee of **2.50 euros**. The entire sum will be paid out to you at the end of the experiment.
- At the end of the experiment, you will be asked to answer a few more questions. Answering these questions completely and honestly is very important for the subsequent evaluation of the experiment. The answers to the questions will naturally be **anonymous** and the questions will only be evaluated for scientific purposes. Your answers in this questionnaire have **NO influence on the payout** in this experiment.

Information about the experiment

- The experiment consists of two parts.
- In the first part you build a team with another participant.
- The participant, who will be on your team, will be randomised. The players in a team will stay **anonymous**. You will never know, who the other member of your team is.
- In this team structure you will play 3 rounds.
- In each round you will solve a team task (see screenshot), for which your team will be rewarded. The reward is therefore a team achievement.
- Your team task is to count together with your team partner how often the digit “7” in the displayed digit block occurs. Every team member counts one digit block. Afterwards enter your result in the input field and **click on Ok.**



- While counting the digit block there is only limited time for you, namely **2 minutes** for every round. A clock shows you how much time is left for you. If you didn't enter the number of 7s after the clock runs out or didn't click OK, the number of 0 will be submitted automatically.
- After you and your team partner have each counted and entered the 7s from the digit block you were working on, the overall result of your team is calculated.
- You have solved the task as a team **successfully**, if
    - The total number of 7s counted by both of you together in the digit blocks deviates from the exact number by a maximum of **four** 7s up or down. Your team receives **10 tokens** as a reward. The 10 tokens are divided up in your team so that you and your team partner receive **5 tokens** each
- You have **not solved** the task as a team, if
    - The total number of 7s counted by both of you together in the digit blocks deviates **by more than four** 7s from the exact number upwards or downwards. In this case, your team will not receive a reward for this round.
- After 3 rounds, your team will be given an additional budget of 5 tokens to donate to the charity organisation *GoAhead!.* Your team doesn't have to contribute to the collection of this donation, so nothing will be deducted from your pay for the donation.
- The aid project receiving the donations is real. It is a *GoAhead!* project, which will build a learning centre for HIV-infected orphans in KwaZulu-Natal in South Africa. You will find additional information at the end of the instructions.
- It is randomly determined which of the two team members is responsible for transferring the **donation budget of 5 tokens** to the charity organisation.
- At this point in the experiment, you are randomly assigned one of two roles: "Player A" or
  "Player B".

Role of Player A

- If you are in the role of Player A, it is in your responsibility to transfer the amount of 5 tokens as a real donation to the charity organisation *GoAhead!.*
- There are two options as to what you can do: You can transfer the 5 tokens to the organisation *GoAhead!* as a real donation. Or you don't transfer part of the money.
- Your decision on the donation budget remains **anonymous**. Your team partner receives information about the decision on the computer screen.
- If you decide to transfer the donation budget to *GoAhead!,* the money will be donated. At the end of the experiment, you can make sure that the amount donated is correct. It will be ensured that the donation goes to *GoAhead!.*
- If you decide not to transfer part of the money, 2 tokens will be added to your existing income. *GoAhead!* will then only receive a donation of 3 tokens. Player B's income remains unchanged.

- [Only Treatment 2: Before you make the decision, you receive the information which alternative action the majority of participants in the role of Player A considered to be the personally morally correct behaviour in previous sessions of the experiment.]
- Consider, that the decision you make here could have an influence on the second part of the experiment:
    - A randomly chosen new team partner can decide to expel you from the team task in part 2, depending on the decision you have made here.
- After Players A and B have made their decision, you receive two assessment questions. For the correct assessment, you can earn additional money.

Role of Player B

As Player B you decide, how you want to behave if Player A does not transfer the donation budget. You have the following possibilities:

- **Ignoring the behaviour of Player A**: You can ignore Player A‘s behaviour. In this case Player A keeps the money, which Player A did not transfer and *GoAhead!* receives the reduced donation.
- **Reporting the behaviour of Player A**: You can report Player A’s behaviour. In this case the whole donation will be transferred to *GoAhead!*. Additionally, an amount of 2 tokens will be deducted from Player A and an amount of 1 token will be deducted from you.
- You make the decision **in parallel** with Player A, i.e., at the time of your decision you do not know whether Player A donates the money or not. You will receive this information after you have made decision.
- If Player A decides to withhold part of the donation budget, your decision will be implemented, and the resulting consequences will be initiated accordingly.
- If Player A decides to pass on the donation budget, each team member will receive the usual period payout from the team task, regardless of what you have decided.
- [Only Treatment 2: Before you make the decision, you receive the information which alternative action the majority of participants in the role of Player B considered to be the personally morally correct behaviour in previous sessions of the experiment.]
- Note that the decision you make here could have an impact on the second part of the experiment:
    - A randomly selected new team partner may exclude you from the team task in part 2 depending on the decision you make here.
- After Players A and B have made their decision, you will receive two assessment questions. You can earn more money for making the correct judgement.

After the two assessment questions, **part 2** of the experiment begins.

- Part 2 of the experiment consists of one round.

- In the second part of the experiment, you will randomly build a new team with another participant who was previously in the same role as you.
- You will either be assigned the role of "Player 1" or "Player 2".
- Player 1 decides whether Player 1 wants to build a team with Player 2 depending on how Player 2 behaved in part 1 of the experiment. This means specifically that player 1 makes their decision for both possible actions of Player 2 in the first part of the experiment.
- Player 1 then receives the information, which decision Player 2 made in the first part of the experiment and the corresponding decision of Player 1 will be implemented.
    - If Player 1 decides to build a team with Player 2, they will work together at the same task as in part 1 and both players will receive 5 tokens each for solving the task correctly.
    - If Player 1 decides not to build a team with Player 2, only Player 1 can earn more tokens. Player 1 will complete the same task as in part 1, but alone on their own. The task is performed correctly if the reported number of 7s in the digit block deviates from the exact number by a maximum of two 7s up or down. Player 1 receives 4 tokens for correctly solving the individual task. Player 2 has no opportunity to earn more tokens and waits until Player 1 has solved the task.
- The experiment is then finished, and you will receive information about the amount you have earned in the experiment and the total amount donated to *GoAhead!* by all teams in this session of the experiment. You will also be informed whether you have answered the assessment question in part 1 of the experiment correctly and thus receive the additional payout.

Payout

- The amount you have earned will be converted at an exchange rate of 1 token = **50 euro cents**
- The amount you have earned will be paid out to you together with a show-up fee of **2.50 euros** after completing the questionnaire.
- The individual transfer amounts to *GoAhead!* will be displayed as a total sum in euros on the results page at the end of the experiment. The money will be transferred to the *GoAhead!* account (account number 6662412, bank code 69490000) by bank transfer. You will find the transfer confirmation on the BaER-Lab homepage one week after the experiment. This is your guarantee that the money has been received by *GoAhead!* and will be used to support the Learning Centre in KwaZulu-Natal.

Please take note

- No communication is allowed during the entire experiment.
- Mobile phones must be switched off for the entire duration of the experiment.

- All decisions you make during the experiment will be anonymous, i.e., none of the other participants will know the identity of the person who has made a particular decision.
- The payout is also anonymous, i.e., none of the participants will find out how much another participant has been paid.

Good luck and thank you for taking part in our experiment!

Additional information *GoAhead!*

GoAhead! is a charitable organisation, founded and run by students from Germany. The learning centre in KwaZulu-Natal (South Africa) is an initiative of *GoAhead!*. It provides HIV-infected orphans with a safe learning environment and warm meals. In addition, the children are offered workshops and sports activities and receive psychological counselling.

You can find more information about *GoAhead!*, the aid project and the donation account at
http://www.goahead-organisation.de.
http://www.goahead-organisation.de/content/helfen/sk/spendenkonto/index_ger.html

You will find confirmation of the donation one week after the experiment at:
http://www.baer-lab.de

Appendix C: Analysis Player A

Table A1: Descriptive data of Players A in Treatments 1 and 2

Treatment	Treatment 1	Treatment 2
n	98	100
Age (average)	23.09	22.14
Gender		
Female	61 (63.54%)	55 (55.560%)
Male	34 (35.42%)	44 (44.44%)*
Non-binary	1 (1.04%)**	
Field of Study		
Business & Economics	53 (55.21%)	65 (65.00%)
Cultural Sciences	34 (35.42%)	29 (29.00%)
Natural Sciences	7 (7.294%)****	3 (3.00%)***
Performance in real-effort task (average payoff in tokens)		
Part 1	12.60	13.35
Part 2	4.18	4.34
Number of exclusions	11 (22.45%)	6 (12.00%)
Personal normative belief		
in support of donation	85 (86.73%)	87 (874%)
in support of embezzlement	13 (13.27%)	13 (13%)
Normative expectation		
in support of donation	80 (81.63%)	
in support of embezzlement	18 (18.37%)	
Empirical expectation		
in support of donation	66 (67.36%)	79 (79%)
in support of embezzlement	32 (32.653%)	21 (21%)
Behaviour		
Donation	78 (79.59%)	85 (85%)
Embezzlement	20 (20.41%)	15 (15%)
Social comparison orientation (average score)	3.409**	3.3085*

*1 answer is missing, **2 answers are missing, ***3 answers are missing, ****4 answers are missing

In total, 198 participants were assigned the role of Player A, of which 98 in Treatment 1, and 100 in Treatment 2. The composition of the treatments differs with regard to age. In Treatment 1, participants are significantly older (average age 23.09 years) than in Treatment 2 (average age 22.14 years; Mann-Whitney U Test: $z = 1.973$, $p = 0.0485$). No significant difference is revealed between genders (Chi-Square Test: $\chi^2 = 2.5468$, $p = 0.280$), field of study ($\chi^2 = 3.3369$, $p = 0.3439$), social comparison orientation ($z = 1.027$, $p = 0.3046$), real-effort task for the first part ($z = -1.408$, $p = 0.1591$), and for the second part of the experiment ($z = -1.044$, $p = 0.2965$). In Treatment 1, eleven subjects were excluded from the task in the second part. In Treatment 2, six subjects were excluded. In Treatment 1, eight of the eleven excluded subjects had embezzled the donation in part I, and in Treatment 2, two of the excluded players had embezzled the donation, the other four had forwarded it.

Discrepancy between personal normative beliefs and normative expectations in Treatment 1: Most participants (n=75; 76.53%) in the role of Player A have been consistent in their personal normative beliefs and normative expectations in support of the donation, compared with only eight Players A (8.16%) who have shown consistent support for the embezzlement of the donation budget. Accordingly, 15 Players A (15.30%) show a discrepancy: 10 (10.20%) hold a personal normative belief in support of the donation but perceive that their peers in the experiment have personal normative expectations in support of the embezzlement. Five (5.10%) hold personal normative beliefs in support of the embezzlement but perceive their peers' personal normative beliefs to be in support of the donation. A McNemar Test shows that there is a slightly significant difference between the ratings of subjects' personal normative beliefs and others' personal normative beliefs (McNemar's $\chi^2 = 2.78$, $p = 0.0956$). Accordingly, only 15 Players A (15.30%) **misperceive** the majority's personal normative beliefs. Thus, the vast majority of subjects in the role of Player A correctly perceive the predominant attitude towards the donation (i.e., not to embezzle it).

Personal normative beliefs, normative expectations, and behaviour (Treatment 1): Of the 85 participants in the role of Player A who believe that donating the money is the right course of action, 76 (89.41%) actually take the corresponding action, while 9 (10.58%) embezzle the donation budget. Of the 13 subjects who personally believe that embezzling the money is right, 11 (84.62%) go on to embezzle the donation budget while two subjects (15.38%) donate it. This results in a statistically significant difference in the donation behaviour between subjects who personally support the donation and those who support the embezzlement (one-sided Chi-Square Test: $\chi^2 = 38.0405$, $p < 0.001$), indicating a positive relationship between personal normative belief and actual behaviour.

We also reveal a significantly positive relationship between normative expectations regarding the donation and actual donation behaviour ($\chi^2 = 29.0482$, $p < 0.001$): Of the 80 subjects who expect their peers to approve of donating, 72 (90.00%) actually donate, while only eight (10.00%) embezzle the donation. Of the 18 Players A who perceive the majority of others to favour embezzling the money, 12 (66.67%) go on to embezzle the donation, and 6 (33.33%) donate the money. This relationship is also evident when we consider only those subjects who personally approve of the donation ($\chi^2 = 18.5948$, $p < 0.001$): Seventy-one subjects (94.67%) who both personally approve of

donating and believe others do so too, actually opt for donating, while only four subjects (5.33%) opt for embezzling. Of those who personally approve of donating but think that most others do not, five subjects (50%) donate while 5 (50%) keep part of the donation budget for themselves.

Taken together, even though we reveal relationships between personal normative beliefs, normative expectations, and donation/embezzlement behaviour, we do not find that participants substantially misperceive the social norm in this context: Only a clear minority of 18.06% misperceives the predominant attitude towards donating the budget, and only a few subjects (18.06%) show a discrepancy between their personal normative beliefs and normative expectations.

Social Information Intervention: Given the already high level of subjects in the role of Player A who donate the money in Treatment 1 (79.59%), we do not find a significant increase in donation behaviour in Treatment 2 (85%), where participants receive the social information intervention (i.e., the information that most others believe that donating the money is the right option) ($\chi^2 = 9948$, $p = 0.319$).

Appendix Chapter 6 – Multiple Normative Expectations and Social Norm Interventions: Experimental Evidence on Whistleblowing Behaviour

Appendix A: Analysis Player B of Deprecated Treatments T0 – T3

In this appendix, we provide information about the analysis of our preliminary treatments that were discontinued due to unforeseen problems in the timing of belief elicitation. These treatments differ from those conducted in our main experiment, as described in our manuscript, in one respect: personal normative beliefs and normative expectations were elicited before the decision of whether or not to transfer the donation budget (Player A) and whether or not to blow the whistle (Player B). In this document, we only focus on Player B. Treatment T0 is the baseline experiment, and T1-T3 are the treatments that involve a social norm intervention (T1 contains a message about whistleblowing, T2 a message about staying silent, T3 a message about both behavioural options). The treatments were conducted between May and June 2024 via the Prolific platform with a total of 444 participants across treatments T0-T3.

In Treatments T0-T3, we excluded 12 participants from the analysis because they failed the attention check. This results in a sample of 91 participants in Treatment T0, 37 in Treatment T1, 41 in Treatment T2, and 41 in Treatment T3. Table 1 provides an overview of our sample regarding socio-demographic characteristics, subjects' behaviour in the experiment, their personal normative beliefs and their normative expectations regarding whistleblowing and staying silent. Randomization successfully achieved a balance across treatments in terms of age (Kruskal-Wallis Test: $\chi2(3) = 0.860$, $p = 0.8351$), gender (Chi-Square Test: $\chi2(6) = 3.9359$, $p = 0.685$), education ($\chi2(18) = 15.8111$, $p = 0.606$), and job position ($\chi2(36) = 41.2387$, $p = 0.252$).

In all treatments, the majority's personal normative belief regarding whistleblowing is that it is either somewhat appropriate or very appropriate (82.42% in T0, 72.97% in T1, 78.05% in T2, and 82.93% in T3) and staying silent is considered as either somewhat inappropriate or very inappropriate (72.53% in T0, 64.86% in T1, 73.17% in T2, and 70.73% in T3). This is largely in line with normative expectations. Most participants perceive that the majority of other participants find whistleblowing somewhat or very appropriate (86.81% in T0, 86.49% in T1, 87.80 in T2, and 82.93% in T3) and staying silent somewhat or very inappropriate (78.02% in T0, 72.97% in T1,

78.05% in T2, and 70.73% in T3). The consistency between personal beliefs and normative expectations suggests that participants predicted the majority's view fairly accurately.

Table 1: Descriptive Statistics of Participants in the Role of Player B across all Treatments T0-T3

	Treatment 0	Treatment 1	Treatment 2	Treatment 3
Age (mean)	34.60	34.86	34.71	34.27
Gender				
Female	47 (48%)	19 (51.35%)	26 (63.41%)	23 (56.10%)
Male	41 (46%)	17 (45.95%)	15 (36.59%)	18 (43.90%)
Non-binary	3 (6%)	1 (2.7%)	-	-
Education				
Student in full-time education	-	-	-	-
School leavers without qualification	1 (1.10%)	-	1 (2.44%)	-
GCSE Level	6 (6.59%)	7 (18.92%)	4 (9.76%)	2 (4.88%)
Completed apprenticeship	6 (6.59%)	3 (8.11%)	5 (12.20%)	6 (14.63%)
A-Level	10 (10.99%)	3 (8.11%)	4 (9.76%)	4 (9.76%)
Undergraduate degree	47 (51.65%)	18 (48.65%)	17 (41.46%)	22 (53.66%)
Postgraduate degree	19 (20.88%)	6 (16.22%)	10 (24.39%)	5 (12.20%)
PhD	2 (2.20%)	-	-	2 (4.88%)
Job position				
Upper management	6 (6.59%)	-	2 (4.88%)	2 (4.88%)
Trained professional	20 (21.98%)	6 (16.22%)	6 (14.63%)	4 (9.76%)
Middle management	19 (20.88%)	8 (21.62%)	7 (17.07%)	14 (34.15%)
Skilled labourer	6 (6.59%)	3 (8.11%)	2 (4.88%)	2 (4.88%)
Junior management	15 (16.48%)	4 (10.81%)	3 (7.32%)	5 (12.20%)
Consultant	5 (5.49%)	-	2 (4.88%)	1 (2.44%)
Administrative staff	8 (8.79%)	8 (21.62%)	12 (29.27%)	5 (12.20%)
Temporary employee	2 (2.20%)	-	-	2 (4.88%)
Support staff	2 (2.20%)	4 (10.81%)	2 (4.88%)	2 (4.88%)
Researcher	1 (1.10%)	1 (2.70%)	-	1 (2.44%)
Student	1 (1.10%)	-	-	-
Self-employed/Partner	6 (6.59%)	2 (5.41%)	2 (4.88%)	3 (7.32%)
Other	-	1 (2.70%)	3 (7.32%)	-
Behaviour				
Silence	31 (34.07%)	12 (32%)	16 (39%)	14 (34%)
Whistleblowing	60 (65.93%)	25 (68%)	25 (61%)	27 (66%)
PNB re whistleblowing				
Very inappropriate	6 (6.59%)	2 (5.41%)	2 (4.88%)	1 (2.44%)
Somewhat inappropriate	10 (10.99%)	8 (21.62%)	7 (17.07%)	6 (14.63%)
Somewhat appropriate	38 (41.76%)	13 (35.14%)	18 (43.90%)	16 (39.02%)
Very appropriate	37 (40.66%)	14 (37.84%)	14 (34.15%)	18 (43.90%)
PNB re silence				
Very inappropriate	26 (28.57%)	10 (27.03%)	16 (39.02%)	15 (36.59%)
Somewhat inappropriate	40 (43.96%)	14 (37.84%)	14 (34.15%)	14 (34.15%)
Somewhat appropriate	20 (21.98%)	10 (27.03)	10 (24.39%)	12 (29.27%)
Very appropriate	5 (5.49%)	3 (8.11%)	1 (2.44%)	-
NE re whistleblowing				
Very inappropriate	5 (5.49%)	1 (2.70%)	-	1 (2.44%)
Somewhat inappropriate	7 (7.69%)	4 (10.81%)	5 (12.20%)	6 (14.63%)
Somewhat appropriate	48 (52.75%)	14 (37.84%)	25 (60.98%)	16 (39.02%)
Very appropriate	31 (34.07%)	18 (48.65%)	11 (26.83%)	18 (43.90%)

NE re silence				
Very inappropriate	22 (24.18%)	11 (29.73%)	18 (43.90%)	14 (34.15%)
Somewhat inappropriate	49 (53.85%)	16 (43.24%)	14 (34.25%)	15 (36.59%)
Somewhat appropriate	17 (18.68%)	8 (21.62%)	6 (14.63%)	11 (26.93%)
Very appropriate	3 (3.30%)	2 (2.37%)	3 (7.32%)	1 (2.44%)
Total	91	37	41	41

Note: PNB = Personal normative beliefs; NE = Normative expectations

Table 2 illustrates the extent to which participants have multiple normative expectations. We observe that, in all four treatments, when participants are asked about their normative expectation before they are making the decision, the vast majority (72.53% in T0, 67.57% in T1, 78.05% in T2, and 68.29% in T3) believe that others view whistleblowing as appropriate and at the same time staying silent as inappropriate. This indicates clear support for the social norm in favour of whistleblowing. Between 9.76% and 18.92% of participants hold the normative expectation that both behavioural options are considered appropriate, and less than 6% across all treatments expect that both options are viewed as inappropriate. Thus, compared to T(Base) in our main experiment, the perception of participants about multiple norms is lower in treatments T0-T3.

Table 2: Distribution of Normative Expectations of Whistleblowing and Silence

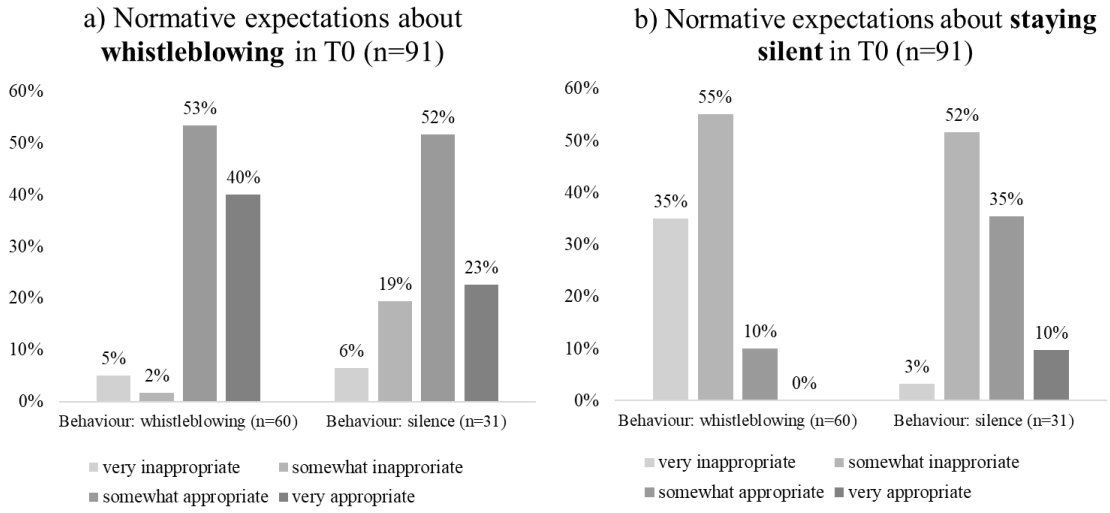
	Treatment T0	**Treatment T1**	**Treatment T2**	**Treatment T3**	**T(Base)**
Normative expectation that both whistleblowing and staying silent are inappropriate	5 (5.49%)	2 (5.41%)	0 (0%)	1 (2.44%)	5 (5.49%)
Normative expectation that whistleblowing is appropriate, and staying silent is inappropriate	66 (72.53%)	25 (67.57%)	32 (78.05%)	28 (68.29%)	55 (60.44%)
Normative expectation that whistleblowing is inappropriate, and staying silent is appropriate	7 (7.69%%)	3 (8.11%)	5 (12.20%)	6 (14.63%)	10 (10.99%)
Normative expectation that both whistleblowing and staying silent are appropriate	13 (14.29%)	7 (18.92%)	4 (9.76%)	6 (14.63%)	21 (23.08%)

In the next step, based on the data of Treatments T0-T3, we will test the predictions that we analysed in the main manuscript for Treatments T(Base), T(WM), T(SM), and T(WM+SM).

The first three predictions are investigated based on the sample of 91 Players B in Treatment T0. Of these 91 subjects, 60 (65.93%) decided to blow the whistle, while 31 (34.07%) stayed silent. As in the primary analysis, we assign values from 1 to 4 to the scale for measuring the appropriateness of whistleblowing and staying silent (i.e., 1 = very inappropriate, 2 = somewhat inappropriate, 3 = somewhat appropriate, 4 = very appropriate). In Treatment T0, the mean value for the normative expectation about the appropriateness of whistleblowing is 3.15 (sd = 0.7877), and for the normative expectation regarding the appropriateness of silence the mean value is 2.01 (sd = 0.7527).

Normative expectations and whistleblowing behaviour: Our first prediction was that the normative expectations regarding the appropriateness of whistleblowing are positively related to whistleblowing behaviour. From our results of Treatment T0, we observe that for those who blow the whistle, the mean value for the appropriateness of whistleblowing is 3.28, while it is 2.90 for those who stay silent. The detailed distribution of the appropriateness ratings regarding whistleblowing can be seen in Figure 1a) for the respective behavioural choices. Comparing the two distributions, a one-sided Mann-Whitney U Test reveals that the appropriateness ratings are significantly higher for whistleblowers than for subjects who stayed silent (z = 2.322, p = 0.010). Thus, in accordance with T(Base) in the main experiment, there is a significantly positive relationship between the normative expectations regarding the appropriateness of whistleblowing and actual whistleblowing behaviour.

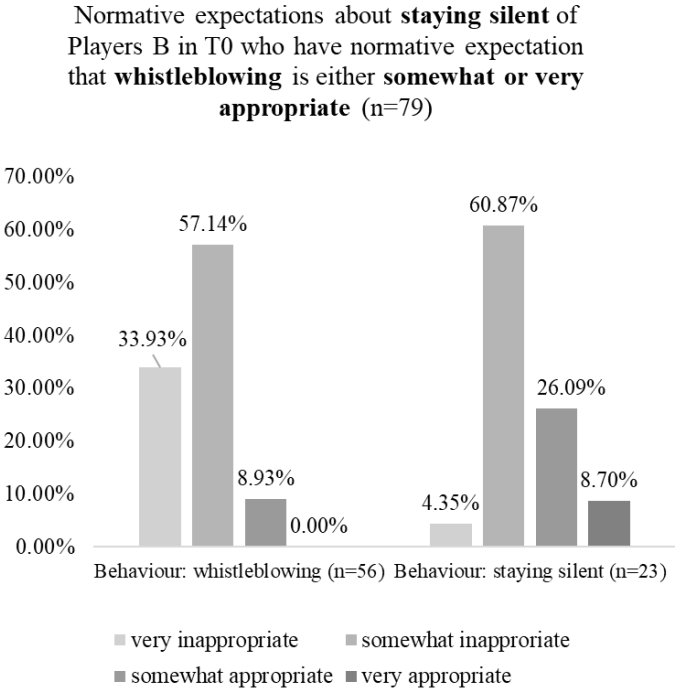
Figure 2: Distribution of Normative Expectations and Behaviour



The second prediction stated that the normative expectations about the appropriateness of staying silent are negatively related to whistleblowing. With a value of 1.75, whistleblowers have a lower mean value for the appropriateness of staying silent compared to subjects who stayed silent, whose mean value is 2.52 in Treatment T0. For the detailed distributions of the appropriateness ratings, we refer to Figure 1b). A one-sided Mann-Whitney U Test reveals that whistleblowers' appropriateness ratings of staying silent are significantly lower than those of subjects who did stay silent ($z = 4.499$, $p < 0.001$). Consequently, based on the data of Treatment T0, our second prediction is also supported in that normative expectations about the appropriateness of staying silent seem to be negatively related to whistleblowing behaviour (in accordance with the analysis of T(Base)).

In the next step, we are interested in how normative expectations regarding the appropriateness of whistleblowing and staying silent are jointly related to the whistleblowing decision. Therefore, we predicted that for individuals with normative expectations that whistleblowing is appropriate (either somewhat appropriate or very appropriate), the normative expectations regarding the appropriateness of staying silent are negatively related to whistleblowing. The investigation of this prediction is based on a sample of 79 subjects in Treatment T0 who have the normative expectation that whistleblowing is either somewhat or very appropriate. In general, these subjects have a mean value of 1.94 for rating the appropriateness of staying silent. The 56 subjects (70.89%) who blew the whistle have a mean value for the appropriateness of staying silent of 1.75, while the 23 subjects (29.11%) who remained silent have a mean value of 2.39. The detailed distribution of appropriateness ratings divided according to the chosen behaviour is displayed in Figure 2. The difference in the distributions of appropriateness ratings of staying silent between whistleblowers and non-whistleblowers is statistically significant (one-sided Mann-Whitney U Test: $z = 3.517$, $p < 0.001$). Thus, for individuals who have normative expectations that whistleblowing is appropriate (either somewhat appropriate or very appropriate), the normative expectations about the appropriateness of staying silent is negatively related to whistleblowing. Summarized, as with T(Base) in the main experiment, all three predictions find full support in this dataset of T0.

Figure 3: Multiple Normative Expectations and Behaviour (T0)



Social norm interventions: The next two predictions refer to the influence of social information interventions. More precisely, we predicted that providing information about the majority's personal normative belief for both behavioural options (i.e., whistleblowing and staying silent) increases whistleblowing behaviour compared to communicating only the majority's personal normative belief for whistleblowing (Prediction 4) or for staying silent (Prediction 5).

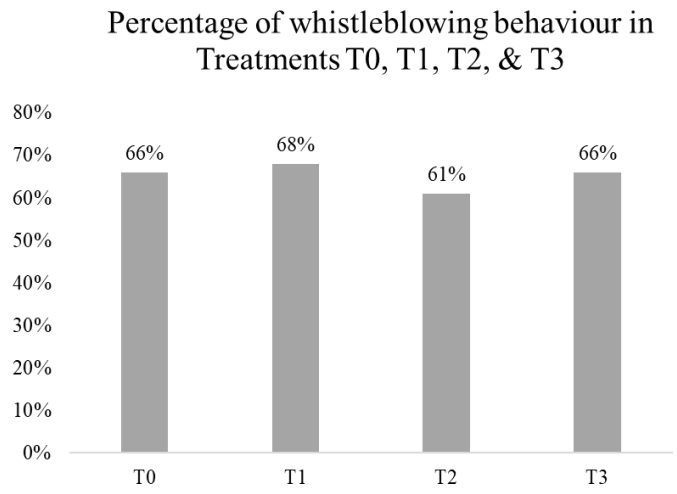
To analyse Prediction 4, we compare the percentage of whistleblowing in Treatment T1 with the share of whistleblowers in Treatment T3. Of the 37 subjects in T1, 25 (67.57%) blew the whistle, while 12 (32.43%) stayed silent. In Treatment T3, the share of whistleblowers is similar: of the 41 subjects, 27 (65.85%) blow the whistle and 14 (34.15%) stay silent. Applying a one-sided Chi-Square Test yields no significant difference in whistleblowing behaviour between the two treatments ($\chi^2(1) = 0.026$, $p = 0.437$). Therefore, unlike the analysis in the primary document between T(WM) and T(WM+SM), we cannot support our fourth prediction based on the data where we elicited personal normative beliefs and normative expectations before the intervention and decision. In this case. providing information about the majority's personal normative belief for both behaviours (i.e., whistleblowing and staying silent) does not increase

whistleblowing behaviour compared to communicating only the majority's personal normative belief about whistleblowing.

To analyse Prediction 5, we compare the share of whistleblowers in Treatment T2 with those in Treatment T3. Of the 41 subjects in Treatment T2, 25 (60.98%) blew the whistle, while 16 (39.02%) stayed silent. Comparing these numbers to the 65.85% whistleblowers in Treatment T3, leads to a statistically insignificant difference between the treatments (one-sided Chi-Square Test: $\chi^2(1) = 0.210$, $p = 0.324$). Consequently, in the setting where we elicit personal normative beliefs and normative expectations before the intervention and decision, we cannot support Prediction 5 either. However, this time, this is in line with the results of our main analysis between T(SM) and T(WM+SM). We conclude that providing information about the majority's personal normative belief for both behavioural options (i.e., whistleblowing and staying silent) does not seem to increase whistleblowing behaviour compared to communicating only its belief about staying silent.

We find no significant differences if we compare the share of whistleblowers in all four treatments where we elicited personal normative beliefs and normative expectations before the intervention and decision (T0-T3). As displayed in Figure 3, the share of whistleblowers is very similar in all four treatments.

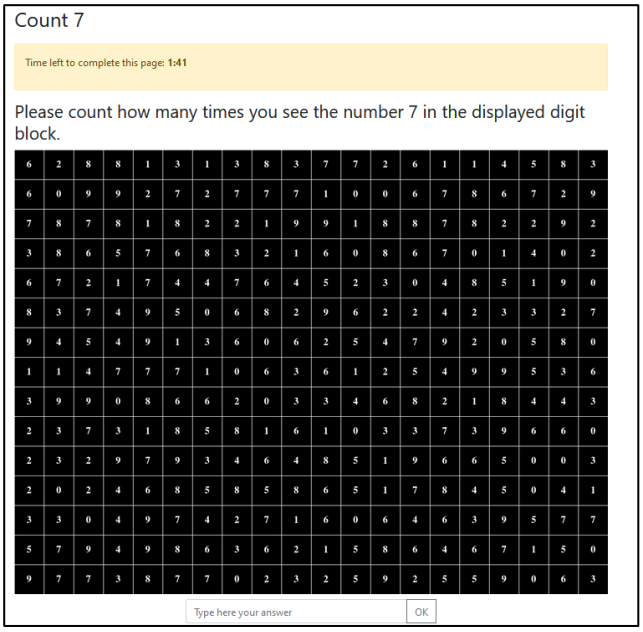
Figure 4: Whistleblowing Behaviour across Treatments



Appendix B: Instructions

Information about the experiment

- The experiment consists of two parts.
- In Part I, you will be randomly assigned to another player, to form a team of two.
- In your team, you will play 2 rounds.
- In each round you will solve a team task, for which your team will get a monetary reward. The success of the task depends on each team member's performance.
- The task is to count how many times you can see the number 7 in the digit block displayed on your screen (see screenshot). Enter your count in the input field and **click OK**.
- Your team member sees a similar digit block and does the same task.



- You have to count and submit your count within a time limit of **2 minutes** – the time is displayed on screen. If you fail to enter your digit count and click OK within the time limit, the number 0 will be submitted automatically.
- Your team **solved** the task **successfully** if the total number of 7s counted by **both of you together** deviates from the exact number by no more than **four digits** up or down (e.g. if the exact number is 10: any number between 6 and 14 would be valid). Your team earns **3 GBP** as a reward, which is divided between the two team members, so you and your team partner receive **1,5 GBP** each.
- You have **not solved** the team task successfully, if your joint total number count deviates **by more than four** from the exact number up or down. In this case, your team will **not** receive any payout for this round.
- After playing 2 rounds, your team will be allocated a budget of **1,5 GBP** dedicated to being donated to the charity *Cancer Research UK*. This donation budget is separate

from your current earnings – **but the decisions you make later on could increase or decrease your earnings.**

- At this point, each team member will be randomly allocated either the role of **Player A** or **Player B**.
- **Player A** is responsible for transferring the donation budget to the charity.
    - If Player A **confirms the transfer**, the full budget will be transferred to the charity. The experimenters make sure that the donation goes to *Cancer Research UK*. Player A's and Player B's earnings remain unaffected.
    - However, Player A can decide to **keep part of the donation budget** for themselves.
- **Player B** has to respond to Player A's decision of keeping parts of the donation budget in one of two ways: either **by overlooking** or **by reporting the behaviour of Player A**.
    - If Player B decides to **overlook** the behaviour of Player A, **1 GBP** of the donation budget will be added to Player A's earnings and the remaining **0,5 GBP** go to *Cancer Research UK*. Player B's earnings remain unaffected.
    - If Player B decides to **report** the behaviour of Player A, Player A has to give back the **1 GBP.** Thus, *Cancer Research UK* receives the full donation budget of 1,5 GBP. In addition, another 1 GBP is deducted from Player A's earnings. Player B will have their payoff deducted by 0,5 GBP.

Decisions		**Changes in earnings**		
Player A	**Player B**	**Donation**	**Player A**	**Player B**
Confirm transfer to donation	Overlook	1,5 GBP	Current earnings **remain unchanged**	
	Report	1,5 GBP	Current earnings **remain unchanged**	
Keep part of donation	Overlook	0,5 GBP	Current earnings **+ 1 GBP**	Current earnings **remain unchanged**
	Report	1,5 GBP	Current earnings **- 1 GBP**	Current earnings **- 0,5 GBP**

Part II of the experiment

- Part II consists of only one round.
- You will be randomly assigned to another participant who was previously in the same role as you. Hence, the team will consist of either two previous Players A or two previous Players B. In the new team, you will be assigned a new role, which we refer to as either the role of "Player 1" or "Player 2".
- Player 1 will be informed about Player 2's choice of action in Part I of the experiment and decides whether they want to form a team with Player 2. Therefore, please note

that the decision you made in Part I potentially impacts the decision that your team partner will make in Part II.

- If Player 1 decides to **accept Player 2 as team member**, they will work together on the same task as in Part I and both players will receive **1,5 GBP** each for solving the task correctly (same as in Part I).
- If Player 1 decides **not to build a team with Player 2**, only Player 1 can earn further money. Player 1 will complete the same task as in Part I, but the task is solved successfully only if the reported number of 7s in the digit block deviates from the exact number by no more than **two 7s** up or down. Player 1 receives only a reduced amount of **1,2 GBP** for correctly solving the individual task. Player 2 cannot make any money and has to wait until Player 1 has solved the task.

- This concludes the experiment, and you will receive information about the amount you earned in the experiment and the amount that has been donated to *Cancer Research UK.*
- Subsequently, we ask you to fill out **a short questionnaire**. This is **mandatory** but your answers do not affect your bonus payments.

Decision of Player 1	Changes in earnings	
	Player 1	Player 2
Accept Player 2	Current earnings **+ 1,5 GBP if solved correctly**	Current earnings **+ 1,5 GBP if solved correctly**
Not accept Player 2	Current earnings **+ 1,2 GBP if solved correctly**	Current earnings **remain unchanged**

Your Payment

- There is a flat payment of **4 GBP** for participating in this experiment.
- Depending on the performance in the team task and your decisions you will receive bonus payments as explained above.
- Your total bonus payment will be displayed to you at the end of the experiment.

Transferring the money to the charity

- The total sum of donations of all the teams that participated in the experiment will be transferred to *Cancer Research UK* (account number 22994289, bank code 56-00-13) by bank transfer. You can find confirmation of the transfer on the following homepage (https://t1p.de/4jrqp) one week after the experiment. In this way you can be assured that the money has been transferred to *Cancer Research UK*. We will message you via Prolific as soon as the donation confirmation has been uploaded onto the homepage.

Thank you for taking part in our experiment!

Appendix C: Analysis of Player A in Main Treatments

Table 1: Descriptive Data of Players A across Treatments

	T(Base)	T(WM)	T(SM)	T(WM+SM)
Total (n)	92	93	94	97
Age (mean)	34.67	37.42	34.12	35.02
Gender				
Female	58 (63.04%)	48 (51.61%)	44 (46.81%)	52 (53.61%)
Male	34 (36.96%)	44 (47.31%)	49 (52.13%)	45 (46.39%)
Non-binary	-	1 (1.08%)	1 (1.06%)	-
Education				
Student in full-time education	1 (1.09%)	4 (4.30%)	-	1 (1.03%)
School leavers without quali.	-	3 (3.23%)	1 (1.06%)	2 (2.06%)
GCSE Level	8 (8.70%)	10 (10.75%)	11 (11.70%)	6 (6.19%)
Completed apprenticeship	1 (1.09%)	7 (7.53%)	4 (4.26%)	3 (3.09%)
A-Level	12 (13.04%)	5 (5.38%)	14 (14.89%)	7 (7.22%)
Undergraduate degree	56 (60.87%)	36 (38.71%)	39 (41.49%)	47 (48.45%)
Postgraduate degree	12 (13.04%)	27 (29.03%)	21 (22.34%)	29 (29.90%)
PhD	2 (2.17%)	1 (1.08%)	4 (4.26%)	13 (2.19%)
Job position				
Upper management	2 (2.17%)	10 (10.75%)	4 (4.26%)	5 (5.15%)
Trained professional	17 (18.48%)	15 (16.13%)	21 (22.34%)	18 (18.56%)
Middle management	21 (22.83%)	23 (24.73%)	22 (23.40%)	21 (21.65%)
Skilled labourer	3 (3.26%)	3 (3.26%)	7 (7.45%)	5 (5.15%)
Junior management	10 (10.87%)	8 (8.60%)	8 (8.51%)	11 (11.34%)
Consultant	2 (2.17%)	2 (2.15%)	3 (3.19%)	1 (1.03%)
Administrative staff	14 (15.22%)	9 (9.68%)	12 (12.77%)	22 (22.68%)
Temporary employee	-	-	-	2 (2.06%)
Support staff	7 (7.61%)	11 (11.83%)	6 (6.38%)	6 (6.19%)
Researcher	2 (2.17%)	1 (1.08%)	-	1 (1.03%)
Student	3 (3.26%)	-	4 (4.26%)	-
Self-employed/Partner	3 (3.26%)	7 (7.53%)	4 (4.26%)	4 (4.12%)
Other	8 (8.70%)	4 (4.30%)	3 (3.19%)	1 (1.03%)
Behaviour				
Embezzlement	17 (18.48%)	18 (19.35%)	17 (18.09%)	16 (16.49%)
Donation	75 (81.52%)	75 (89.65%)	77 (81.91%)	81 (83.50%)
PNB re Embezzlement				
Very inappropriate	42 (45.57%)	36 (38.71%)	48 (51.06%)	41 (42.27%)
Somewhat inappropriate	33 (35.87%)	37 (39.78%)	23 (24.47%)	35 (36.08%)
Somewhat appropriate	12 (13.04%)	16 (17.20%)	16 (17.02%)	15 (15.46%)
Very appropriate	5 (5.43%)	4 (4.30%)	7 (7.45%)	6 (6.19%)
PNB re Donation				
Very inappropriate	-	4 (4.30%)	4 (4.26%)	1 (1.03%)
Somewhat inappropriate	1 (1.09%)	2 (2.15%)	5 (5.32%)	5 (5.15%)
Somewhat appropriate	14 (15.22%)	21 (22.58%)	16 (17.02%)	18 (18.56%)
Very appropriate	77 (83.70%)	66 (90.97%)	69 (73.40%)	73 (75.26%)
NE re Embezzlement				
Very inappropriate	35 (38.04%)			
Somewhat inappropriate	43 (46.74%)			
Somewhat appropriate	10 (10.87%)			
Very appropriate	4 (4.35%)			
NE re Donation				
Very inappropriate	-			
Somewhat inappropriate	1 (1.01%)			
Somewhat appropriate	25 (27.17%)			
Very appropriate	66 (71.74%)			

Note: PNB = Personal normative belief; NE = normative expectations

We had to exclude 12 participants in the role of Player A from the analysis because they failed the attention check included in the questionnaire at the end of the experiment. This results in a total sample of 376 participants in the role of Player A: 92 in T(Base), 93 in T(WM), 94 in T(SM) and 97 in T(WM+SM). Across the treatments (Table 1), there is no significant difference with respect to the variables gender (Chi-Square Test: $\chi^2(6) = 6.1787$, $p = 0.333$) and job position ($\chi^2(36) = 44.3013$, $p = 0.161$). However, there is a difference in age (Kruskal Wallis Test $\chi^2(3) = 7.348$ $p = 0.0616$), especially in T(WM) the average age is higher compared to other treatments. Moreover, the education level differs between treatments ($\chi^2(21) = 37.1167$, $p = 0.016$).

We treat each participant as one independent unit of observation. Across all treatments more than 80% of the participants donated the money and the majority holds the personal normative belief that donating the full amount to the charity is very appropriate (between 73.40% and 90.97%). Similarly, the majority personal normative belief in each treatment is that embezzling the money is somewhat or very inappropriate (> 75%).

Multiple Norms: As can be seen from Table 2 the majority of participants in T(Base) has consistent personal normative beliefs that donation is appropriate, and embezzlement is inappropriate (80.43%). About 20% hold multiple, inconsistent personal normative beliefs, especially those who find both behavioural options appropriate. Concerning normative expectations, we observe a similar pattern, with more than 80% assuming others hold the belief that making the donation in full is appropriate, and embezzling the money inappropriate (Table 3). Moreover, 13 out of 92 participants expect others to find both behavioural options appropriate.

Table 2: Personal Normative Beliefs in T(Base)

		Personal normative belief: Donation		
		Inappropriate	Appropriate	Σ
Personal normative belief: Embezzlement	Inappropriate	1 (1.09%)	74 (80.43%)	75
	Appropriate	0 (0%)	17 (18.48%)	17
	Σ	1	92	92

Table 3: Normative Expectations in T(Base)

		Normative expectation: Donation		
		Inappropriate	Appropriate	Σ
Normative expectation: Embezzlement	Inappropriate	0 (0%)	78 (84.78%)	78
	Appropriate	1 (1.09%)	13 (14.13%)	14
	Σ	1	91	92

Discrepancy between personal normative beliefs and normative expectations: Within subjects we do not find any significant difference between personal normative beliefs and normative expectations in T(Base) regarding embezzling the budget (Wilcoxon signed-rank test: $z = 0.360$, $p = 0.7191$). However, regarding donating the full budget we find a significant difference between personal normative beliefs and normative expectations (Wilcoxon signed-rank test: $z = 2.294$, $p = 0.0218$). The mean personal normative beliefs for donation appropriateness is 3.83 (sd=0.04) compared to the mean normative expectations for donation appropriateness of 3.71 (sd = 0.05). This means that a substantial share of individuals believed that others view donating the full amount as less appropriate than they do.

Influence of the Social Information Intervention: For the sake of completeness, we analysed the influence of a social information intervention on Players' A donation/embezzlement decision. When participants received the information about the majority's norm concerning appropriate donation behaviour from the baseline treatment, donation behaviour in T(WM) did ´not significantly change compared to T(Base) (two-sided test, $\chi^2(1) = 0.0232$, $p = 0.879$). We also find no significant difference when participants are provided with the information of the majority norm concerning embezzlement T(SM) compared to T(Base), ($\chi^2(1) = 0.0048$, $p = 0.945$), or with both information T(WM+SM) compared to T(Base), ($\chi^2(1) = 0.1289$, $p = 0.720$). As the donation rate is already above 80% in the baseline T(Base) it is not surprising that neither of the interventions significantly increased the donation behaviour.

Appendix D: Tables

Table A1: Personal Normative Beliefs in T(Base)

		Personal normative belief: Silence		
		Inappropriate	Appropriate	Σ
Personal normative belief: Whistleblowing	Inappropriate	7 (7.69%)	17 (18.68%)	24 (26.37%)
	Appropriate	50 (54.95%)	17 (16.68%)	67 (71.63%)
	Σ	57 (62.64%)	34 (35.36%)	91

Table A2: Personal Normative Beliefs and Behaviour in T(Base). Subsample: participants who personally believe that whistleblowing is appropriate and silent inappropriate vs. both behavioural alternatives are appropriate.

		Personal normative belief		
		Whistleblowing appropriate & silent inappropriate	Both appropriate	Σ
Behaviour	Silent	12 (17.91%)	14 (20.90%)	26 (38.81%)
	Whistleblowing	38 (56.72%)	3 (4.48%)	41 (61.19%)
	Σ	50 (74.63%)	17 (25.37%)	67

Table A3: Normative Expectations and Behaviour in T(Base). Subsample: participants who believe that the majority finds whistleblowing appropriate and silent inappropriate vs. both behavioural alternatives appropriate.

		Normative expectations		
		Whistleblowing appropriate & silent inappropriate	Both appropriate	Σ
Behaviour	Silent	19 (25.00%)	15 (19.74%)	34 (44.74%)
	Whistleblowing	36 (47.37%)	6 (7.89%)	42 (55.26%)
	Σ	55 (72.37%)	21 (27.63%)	76

Appendix Chapter 7 – Testing Variations of the Two-Step Norm Elicitation Procedure

Appendix A: Tables

Table A1: Descriptive Analysis across Treatments

	PTD	PTR	T(Base)	T(Before)	T(No Incentive)	T(Questions)
Total (n)	102	99	111	109	110	108
Age (mean)	40.08	38.69	42.11	40.58	39.76	39.21
Gender						
Female	61 (59.80%)	56 (56.57%)	58 (52.25%)	52 (47.71%)	54 (49.09%)	66 (61.11%)
Male	41 (20.20)	42 (42.42%)	53 (47.75%)	56 (51.38%)	56 (50.91%)	42 (38.89%)
Non-binary	-	1 (1.01%)	-	1 (0.92%)	-	-
Education						
Student in full-time education	-	1 (1.01%)	2 (1.80%)	1 (0.92%)	1 (0.91%)	-
School leavers without qualification	4 (3.92%)	1 (1.01%)	2 (1.80%)	1 (0.92%)	1 (0.91%)	2 (1.85%)
GCSE Level	18 (17.65%)	12 (12.12%)	15 (13.51%)	15 (13.76%)	14 (12.73%)	12 (11.11%)
Completed apprenticeship	1 (0.98%)	5 (5.05%)	5 (3.60%)	5 (4.59%)	6 (5.45%)	5 (4.63%)
A-Level	20 (19.61%)	22 (22.22%)	11 (9.91%)	21 (19.27%)	12 (10.91%)	20 (18.52%)
Undergraduate degree	41 (40.20%)	37 (37.37%)	52 (46.85%)	45 (41.28%)	52 (47.27%)	50 (46.30%)
Postgraduate degree	18 (17.65%)	21 (21.21%)	20 (18.02%)	16 (14.68%)	22 (20%)	14 (12.96%)
PhD	-	-	5 (4.50%)	5 (4.59%)	2 (1.82%)	5 (4.63%)
Employment status						
Yes, full-time	53 (51.96%)	53 (51.96%)	68 (61.26%)	55 (50.46%)	58(52.73%)	60 (55.56%)
Yes, part-time	24.51%)	25 (24.51%)	22 (19.82%)	25 (22.94%)	30 (27.27%)	24 (22.22%)
No	24 (23.53%)	24 (23.53%)	21 (18.92%)	29 (26.61%)	22 (20%)	24 (22.22%)
Role at work						
Upper management	2 (1.96%)	4 (4.04%)	3 (2.70%)	3 (2.75%)	2 (1.82%)	5 (4.63%)
Trained professional	20 (19.61)	12 (12.12%)	13 (11.71%)	21 (19.27%)	13 (11.82%)	17 (15.74%)
Middle management	15 (14.71%)	19 (19.19%)	22 (19.82%)	16 (14.68%)	19 (17.27%)	17 (15.74%)
Skilled labourer	7 (6.86%)	8 (8.08%)	4 (3.60%)	3 (2.75%)	2 (1.82%)	4 (3.70%)
Junior management	9 (8.82%)	8 (8.08%)	17 (15.32%)	6 (5.50%)	9 (8.18%)	4 (3.70%)
Consultant	2 (1.96%)	5 (5.05%)	-	1 (0.92%)	5 (4.55%)	5 (4.63%)
Administrative staff	10 (9.80%)	13 (13.13%)	16 (14.41%)	15 (13.76%)	17 (15.45%)	15 (13.89%)
Temporary employee	2 (1.96%)	-	1 (0.90%)	1 (0.92%)	-	-
Support staff	5 (4.90%)	3 (3.03%)	6 (5.41%)	7 (6.42%)	5 (4.55%)	3 (2.78%)
Researcher	-	1 (1.01%)	1 (0.90%)	-	1 (0.91%)	1 (0.93%)
Student	3 (2.94%)	3 (3.03%)	2 (1.80%)	2 (1.83%)	5 (4.55%)	3 (2.73%)
Self-employed/Partner	7 (6.86%)	6 (6.06%)	7 (6.31%)	9 (8.26%)	12 (10.01%)	8 (7.41%)
Other	3 (2.94%)	5 (5.05%)	3 (2.70%)	5 (4.59%)	3 (2.73%)	4 (3.70%)
Not working	17 (16.67%)	12 (12.12%)	16 (14.41%)	20 (18.35%)	17 (15.45%)	22 (20.37%)
Decision / Share						
A (£2.25 / £0.25)	22 (21.57%)		18 (16.22%)	16 (14.68%)	17 (15.45%)	15 (13.89%)
B (£2.00 / £0.50)	6 (5.88%)		9 (8.11%)	8 (7.34%)	11 (10%)	7 (6.48%)
C (£1.50 / £1.00)	16 (15.69%)		15 (13.51%)	13 (11.93%)	12 (10.91%)	12 (11.11%)
D (£1.25 / £1.25)	57 (55.88%)		67 (60.36%)	72 (66.06%)	67 (60.91%)	74 (68.52%)
E (£1.00 / £1.50)	1 (0.98%)		1 (0.90%)	-	-	-
F (£0.50 / £2.00)	-		-	-	1 (0.91%)	-
G (£0.25 / £2.25)	-		1 (0.90%)	-	2 (1.82%)	-
PNB re fair share						
Very inappropriate	7 (6.86%)	4 (4.04%)	3 (2.70%)	1 (0.92%)	2 (1.82%)	3 (2.78%)
Somewhat inappro.	4 (3.92%)	2 (2.02%)	2 (1.80%)	2 (1.83%)	11 (10%)	2 (1.85%)
Somewhat appropriate	30 (29.41%)	29(29.29%)	37 (33.33%)	40 (36.70%)	31 (28.18%)	45 (41.67%)
Very appropriate	61 (59.80%)	64 (64.95%)	69 (62.16%)	66 (60.55%)	66 (60%)	58 (53.70%)
PNB re low share						
Very inappropriate						44 (40.74%)
Somewhat inappro.						53 (49.07%)
Somewhat appropriate						7 (6.48%)

Very appropriate						4 (3.70%)
NE re fair share						
Very inappropriate	5 (4.90%)	2 (2.02%)	2 (1.80%)	2 (1.83%)	2 (1.82%)	2 (1.85%)
Somewhat inappro.	3 (2.94%)	7 (7.07%)	2 (1.80%)	5 (4.59%)	17 (15.45%)	8 (7.41%)
Somewhat appropriate	53 (51.96%)	48 (48.48%)	62 (55.86)	47 (43.12%)	60 (54.55%)	60 (54.55%)
Very appropriate	41 (40.20%)	42 (42.42%)	45 (40.54%)	55 (50.46%)	31 (28.18%)	38 (35.19%)
NE re low share						
Very inappropriate						22 (20.37%)
Somewhat inappro.						56 (51.85%)
Somewhat appropriate						20 (18.52%)
Very appropriate						10 (9.26%)
EE re fair slit						
yes	65 (63.73%)	64 (64.65%)	70 (63.06%)	76 (69.72%)	69 (62.73%)	80 (74.07%)
no	37 (36.27%)	35 (35.35%)	41 (36.94%)	33 (30.28%)	41 (37.27%)	28 (25.93%)

Note: PTD = Pretest Dictator, PTR = Pretest Receiver, PNB = Personal normative beliefs, NE = Normative expectations, EE = Empirical expectations

Appendix B: Instructions

Instructions Pre-Test

Information about the Experiment

Thank you for participating! You will earn a fixed amount of £1 after completing this study. Additional earnings will be transferred as a bonus payment after the study within a couple of days. The instructions explain how to make decisions. Please read these instructions carefully!

You will be randomly and anonymous paired with another participant. You will never be informed of the identity about this participant. Similarly, your assigned partner will never be informed about your identity. All the decisions will be anonymous.

You will be either assigned to the role of the **Divider** or of the **Receiver**.

Your Task if you are in the role of the Divider

If you are in the role of the Divider, your task is to divide £2.5 between you and your assigned Receiver. The Divider must choose an Option A to G (see the table of dividing options below). The Dividing Option determines how much of £2.5 will go to the Divider (you) and how much will go to the Receiver.

The Divider must choose only one of the options.

Dividing options:

Possible dividing options	The option is
A	Divider gets **£2.25** and Receiver gets **£0.25**
B	Divider gets **£2.00** and Receiver gets **£0.50**
C	Divider gets **£1.50** and Receiver gets **£1.00**
D	Divider gets **£1.25** and Receiver gets **£1.25**
E	Divider gets **£1.00** and Receiver gets **£1.50**
F	Divider gets **£0.50** and Receiver gets **£2.00**
G	Divider gets **£0.25** and Receiver gets **£2.25**

If you are in the role of the Receiver

If you are assigned to the role of Receiver, you have no decision to make. You will receive the decision of your assigned Divider and the according share after the study. You cannot react to the decision.

Further Proceedings

Throughout the study, you will be asked some estimation questions for which you may receive a bonus payment if correctly answered. The study ends with a questionnaire.

Random Assignment of Roles
You will be randomly assigned either to the role of the **Divider** or of the **Receiver**.

- If you are to be the **Divider**, your decision will be implemented. Your assigned Receiver is informed about your decision and will receive the bonus payment accordingly.
- If you are to be the **Receiver**, the decision of your assigned Divider will be implemented. You will receive the bonus payment according to your Dividers decision.

Payment
Your fixed payment will be transferred to your Prolific account directly after finishing the study. The additional earnings will be transferred as a bonus payment after the study within a few days.

Divider and Receiver will always remain anonymously during the study.

End of instructions.

Instructions T0-T3

Information about the Experiment

Thank you for participating! You will earn a fixed amount of £1 after completing this study. Additional earnings will be transferred as a bonus payment after the study within a couple of days. The instructions explain how to make decisions. Please read these instructions carefully!

You will be randomly and anonymous paired with another participant. You will never be informed of the identity about this participant. Similarly, your assigned partner will never be informed about your identity. All the decisions will be anonymous.

You will be either assigned to the role of the **Divider** or of the **Receiver.** In the first place, you are both in the role of the **Divider**. After the study, either yours or your assigned partners decision as Divider will be implemented (randomly chosen) while the other person will become the **Receiver** whose decision as Divider is cancelled (and not part of the bonus payment).

Your Task as Divider

Your task as the Divider is to divide £2.5 between you and your assigned partner.

You must choose a Dividing Option A to G (see the table of dividing options below). The Dividing Option determines how much of £2.5 will go to the Divider (you) and how much will go to the Receiver. You may choose only one of the options.

Dividing options:

Possible dividing options	The option is
A	Divider gets **£2.25** and Receiver gets **£0.25**
B	Divider gets **£2.00** and Receiver gets **£0.50**
C	Divider gets **£1.50** and Receiver gets **£1.00**
D	Divider gets **£1.25** and Receiver gets **£1.25**
E	Divider gets **£1.00** and Receiver gets **£1.50**
F	Divider gets **£0.50** and Receiver gets **£2.00**
G	Divider gets **£0.25** and Receiver gets **£2.25**

Further Proceedings

Throughout the study, you will be asked some estimation questions for which you may receive a bonus payment if correctly answered. The study ends with a questionnaire.

Random Assignment of Roles

After the study, you will be randomly assigned to either the role of the **Divider** or of the **Receiver**.

- If you are randomly chosen to be the **Divider**, your decision will be implemented. Your assigned Receiver is informed about your decision and will receive the bonus payment accordingly.
- If you are randomly chosen to be the **Receiver**, the decision of your assigned Divider will be implemented. You cannot react to the decision. You will receive the bonus payment according to your Dividers decision.

Payment
Your fixed payment will be transferred to your Prolific account directly after finishing the study. The additional earnings will be transferred as a bonus payment after the study within a few days.

Divider and Receiver will always remain anonymously during the study.

End of instructions.

REFERENCES

Abad-Segura, E., Castillo-Díaz, F. J., Batlles-delaFuente, A., & Belmonte-Ureña, L. J. (2024). Enhancing competitiveness and sustainability in Spanish agriculture: The role of technological innovation and corporate social responsibility. *Business Strategy & Development*, *7*(4), e70021. https://doi.org/10.1002/bsd2.70021

Abdallah-Ou-Moussa, S., Wynn, M., Kharbouch, O., & Rouaine, Z. (2024). Digitalization and Corporate Social Responsibility: A Case Study of the Moroccan Auto Insurance Sector. *Administrative Sciences*, *14*(11), 282. https://doi.org/10.3390/admsci14110282

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*(4), 1115–1153.

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization*. *British Journal of Social Psychology*, *29*(2), 97–119. https://doi.org/10.1111/j.2044-8309.1990.tb00892.x

Acciarini, C., Borelli, F., Capo, F., Cappa, F., & Sarrocco, C. (2022). Can digitalization favour the emergence of innovative and sustainable business models? A qualitative exploration in the automotive sector. *Journal of Strategy and Management*, *15*(3), 335–352. https://doi.org/10.1108/JSMA-02-2021-0033

Afe, C. E. I., Abodohoui, A., Mebounou, T. G. C., & Karuranga, E. (2019). Perceived organizational climate and whistleblowing intention in academic organizations: Evidence from Selçuk University (Turkey). *Eurasian Business Review*, *9*(3), 1–20. https://doi.org/10.1007/s40821-018-0110-3

Agafonova, A. N., Yakhneeva, I. V., & Mukhametshina, G. R. (2021). Human-Centric Marketing in the Digital Era. In S. I. Ashmarina, J. Horák, J. Vrbka, & P. Šuleř (Hrsg.), *Economic Systems in the New Era: Stable Systems in an Unstable World* (Bd. 160, S. 10–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-60929-0_2

Agerström, J., Carlsson, R., Nicklasson, L., & Guntell, L. (2016). Using descriptive social norms to increase charitable giving: The power of local norms. *Journal of Economic Psychology*, *52*, 147–153. https://doi.org/10.1016/j.joep.2015.12.007

Agrawal, R., Majumdar, A., Kumar, A., & Luthra, S. (2023). Integration of artificial intelligence in sustainable manufacturing: Current status and future opportunities. *Operations Management Research*, *16*(4), 1720–1741. https://doi.org/10.1007/s12063-023-00383-y

Agrawal, R., Surendra Yadav, V., Majumdar, A., Kumar, A., Luthra, S., & Arturo Garza-Reyes, J. (2023). Opportunities for disruptive digital technologies to ensure circularity in supply Chain: A critical review of drivers, barriers and challenges. *Computers & Industrial Engineering*, *178*, 109140. https://doi.org/10.1016/j.cie.2023.109140

Aguilera, R. V., Rupp, D. E., Williams, C. A., & Ganapathi, J. (2007). Putting the S Back in Corporate Social Responsibility: A Multilevel Theory of Social Change in Organizations. *The Academy of Management Review*, *32*(3), 836–863.

Aguinis, H. (2011). Organizational responsibility: Doing good and doing well. In *APA Handbook of Industrial and Organizational Psychology* (S. 855–879). American Psychological Association.

Aitken, M., Ng, M., Horsfall, D., Coopamootoo, K. P. L., Van Moorsel, A., & Elliott, K. (2021). In pursuit of socially-minded data-intensive innovation in banking: A focus group study of public expectations of digital innovation in banking. *Technology in Society*, *66*, 101666. https://doi.org/10.1016/j.techsoc.2021.101666

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Albert, D. A., & Smilek, D. (2023). Comparing attentional disengagement between Prolific and MTurk samples. *Scientific Reports*, *13*(1), 20574. https://doi.org/10.1038/s41598-023-46048-5

Aldboush, H. H. H., & Ferdous, M. (2023). Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. *International Journal of Financial Studies*, *11*(3), 90. https://doi.org/10.3390/ijfs11030090

Aleksić, D., Černe, M., & Batistič, S. (2024). Understanding meaningful work in the context of technostress, COVID-19, frustration, and corporate social responsibility. *Human Relations*, *77*(3), 426–451. https://doi.org/10.1177/00187267221139776

Alfalah, A. A., Muneer, S., & Hussain, M. (2022). An empirical investigation of firm performance through corporate governance and information technology investment with mediating role of corporate social responsibility: Evidence from Saudi Arabia telecommunication sector. *Frontiers in Psychology*, *13*, 959406. https://doi.org/10.3389/fpsyg.2022.959406

Al-Khatib, A. W. (2023). The impact of industrial Internet of things on sustainable performance: The indirect effect of supply chain visibility. *Business Process Management Journal*, *29*(5), 1607–1629. https://doi.org/10.1108/BPMJ-03-2023-0198

Allal-Chérif, O., Costa Climent, J., & Ulrich Berenguer, K. J. (2023). Born to be sustainable: How to combine strategic disruption, open innovation, and process digitization to create a sustainable business. *Journal of Business Research*, *154*, 113379. https://doi.org/10.1016/j.jbusres.2022.113379

Alleyne, P., Haniffa, R., & Hudaib, M. (2019). Does group cohesion moderate auditors' whistleblowing intentions? In *Journal of International Accounting, Auditing and Taxation* (Bd. 34). Elsevier Inc. https://doi.org/10.1016/j.intaccaudtax.2019.02.004

Alleyne, P., Hudaib, M., & Pike, R. (2013). Towards a conceptual model of whistleblowing intentions among external auditors. *The British Accounting Review*, *45*(1), 10–23.

Almeida, R., Pérez-López, J. Á., & Abreu, R. (2022). Digital Corporate Social Responsibility Reporting in the Water Industry. *Water Resources Management*, *36*(11), 3929–3947. https://doi.org/10.1007/s11269-022-03132-1

Al-Omoush, K., Ribeiro-Navarrete, B., & McDowell, W. C. (2024). The impact of digital corporate social responsibility on social entrepreneurship and organizational resilience. *Management Decision*, *62*(8), 2621–2640. https://doi.org/10.1108/MD-11-2022-1613

Al-Omoush, K. S. (2024). Drivers of digital corporate social responsibility during unprecedented crises: An institutional perspective. *Kybernetes*, *53*(3), 882–900. https://doi.org/10.1108/K-07-2022-0959

Ambrose, M. L., Arnaud, A., & Schminke, M. (2007). Individual Moral Development and Ethical Climate: The Influence of Person–Organization Fit on Job Attitudes. *Journal of Business Ethics*, *77*(3), 323–333. https://doi.org/10.1007/s10551-007-9352-1

Amini, M., & Bienstock, C. C. (2014). Corporate sustainability: An integrative definition and framework to evaluate corporate practice and guide academic research. *Journal of Cleaner Production*, *76*, 12–19. https://doi.org/10.1016/j.jclepro.2014.02.016

Andre, P., Boneva, T., Chopra, F., & Falk, A. (2024). Misperceived Social Norms and Willingness to Act Against Climate Change. *Review of Economics and Statistics*, 1–46.

Andreoni, J., & Bernheim, B. D. (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384

Anvari, F., Wenzel, M., Woodyatt, L., & Haslam, S. A. (2019). The social psychology of whistleblowing: An integrated model. *Organizational Psychology Review*, *9*(1), 41–67. https://doi.org/10.1177/2041386619849085

Ardito, L. (2023). The influence of firm digitalization on sustainable innovation performance and the moderating role of corporate sustainability practices: An empirical investigation. *Business Strategy and the Environment*, *32*(8), 5252–5272. https://doi.org/10.1002/bse.3415

Ardito, L., Raby, S., Albino, V., & Bertoldi, B. (2021). The duality of digital and environmental orientations in the context of SMEs: Implications for innovation performance. *Journal of Business Research*, *123*, 44–56. https://doi.org/10.1016/j.jbusres.2020.09.022

Arnal-Pastor, M., & Berné-Martínez, J.-M. (2024). A semantic analysis of social innovation and corporate social responsibility in the Spanish digital press. *Review of Managerial Science*. https://doi.org/10.1007/s11846-024-00743-3

Aslaksen, H. M., Hildebrandt, C., & Johnsen, H. Chr. G. (2021). The long-term transformation of the concept of CSR: Towards a more comprehensive emphasis on sustainability. *International Journal of Corporate Social Responsibility*, *6*(1), 11. https://doi.org/10.1186/s40991-021-00063-9

Atanasov, A., Chipriyanova, G., & Krasteva-Hristova, R. (2023). Integration of Digital Technologies in Corporate Social Responsibility (CSR) Activities: A Systematic Literature Review and Bibliometric Analysis. *Journal of Risk and Financial Management*, *16*(8), 373. https://doi.org/10.3390/jrfm16080373

Aycinena, D., Bogliacino, F., & Kimbrough, E. O. (2024). Measuring norms: Assessing the threat of social desirability bias to the Bicchieri and Xiao elicitation method. *Journal of Economic Behavior & Organization*, *222*, 225–239. https://doi.org/10.1016/j.jebo.2024.04.002

Ayu Purnamawati, I. G., Yuniarta, G. A., & Jie, F. (2023). Strengthening the role of corporate social responsibility in the dimensions of sustainable village economic development. *Heliyon*, *9*(4), e15115. https://doi.org/10.1016/j.heliyon.2023.e15115

Bag, S., Telukdarie, A., Pretorius, J. H. C., & Gupta, S. (2021). Industry 4.0 and supply chain sustainability: Framework and future research directions. *Benchmarking: An International Journal*, *28*(5), BIJ-03-2018-0056. https://doi.org/10.1108/BIJ-03-2018-0056

Bandura, A. (1971). *Social learning theory*. General Learning Press.

Bansal, P., & Song, H.-C. (2017). Similar But Not the Same: Differentiating Corporate Sustainability from Corporate Responsibility. *Academy of Management Annals*, *11*(1), 105–149. https://doi.org/10.5465/annals.2015.0095

Barkoukis, V., Petrou, M., Lazuras, L., & Ourda, D. (2021). An empirical investigation of sport stakeholders' beliefs about whistleblowing against doping behaviour. *International Journal of Sport and Exercise Psychology*, *0*(0), 1–18. https://doi.org/10.1080/1612197X.2021.1948585

Barnett, M. L., Henriques, I., & Husted, B. W. (2020). The Rise and Stall of Stakeholder Influence: How the Digital Age Limits Social Control. *Academy of Management Perspectives*, *34*(1), 48–64. https://doi.org/10.5465/amp.2017.0080

Barr, A., Lane, T., & Nosenzo, D. (2018). On the social inappropriateness of discrimination. *Journal of Public Economics*, *164*, 153–164. https://doi.org/10.1016/j.jpubeco.2018.06.004

Barth, M., Gossen, M., Lang, D. J., & Santarius, T. (2023). Sustainable digitalization - fostering the twin transformation in a transdisciplinary way. *GAIA - Ecological Perspectives for Science and Society*, *32*(1), 6–9. https://doi.org/10.14512/gaia.32.S1.3

Bartuli, J., Mir Djawadi, B., & Fahr, R. (2016). Business Ethics in Organizations: An Experimental Examination of Whistleblowing and Personality. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2840134

Bašić, Z., & Verrina, E. (2023). Personal norms—And not only social norms—Shape economic behavior. *MPI Collective Goods Discussion Paper*, *2020/25*.

Bednárová, M., & Serpeninova, Y. (2023). Corporate digital responsibility: Bibliometric landscape – chronological literature review. In *International Journal of Digital Accounting Research* (Bd. 23). https://doi.org/10.4192/1577-8517-v23_1

Belhadi, A., Kamble, S., Gunasekaran, A., & Mani, V. (2022). Analyzing the mediating role of organizational ambidexterity and digital business transformation on industry 4.0 capabilities and sustainable supply chain performance. *Supply Chain Management: An International Journal*, *27*(6), 696–711. https://doi.org/10.1108/SCM-04-2021-0152

Benešová, A., Basl, J., Tupa, J., & Steiner, F. (2021). Design of a business readiness model to realise a green industry 4.0 company. *International Journal of Computer Integrated Manufacturing*, *34*(9), 920–932. https://doi.org/10.1080/0951192X.2021.1946858

Bernini, F., Ferretti, P., Gonnella, C., & La Rosa, F. (2024). Measuring machinewashing under the corporate digital responsibility theory: A proposal for a methodological path. In *Business Ethics, the Environment and Responsibility*. https://doi.org/10.1111/beer.12653

Bernstein, E. S. (2017). Making Transparency Transparent: The Evolution of Observation in Management Theory. *Academy of Management Annals*, *11*(1), 217–266. https://doi.org/10.5465/annals.2014.0076

Bhanot, S. P. (2021). Isolating the effect of injunctive norms on conservation behavior: New evidence from a field experiment in California. *Organizational Behavior and Human Decision Processes*, *163*, 30–42.

Bhattacharyya, S. S. (2023). Exploration and explication of the nature of online reviews of organizational corporate social responsibility initiatives. *International Journal of Organizational Analysis*, *31*(6), 2280–2299. https://doi.org/10.1108/IJOA-10-2021-2994

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C. (2017). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, *23*(2), 161–178. https://doi.org/10.1002/bdm.648

Bicchieri, C., & Dimant, E. (2022a). Nudging with care: The risks and benefits of social information. *Public Choice*, *191*(3–4), 443–464.

Bicchieri, C., & Dimant, E. (2022b). Nudging with care: The risks and benefits of social information. *Public Choice*, *191*(3–4), 443–464. https://doi.org/10.1007/s11127-019-00684-6

Bicchieri, C., Dimant, E., Gachter, S., & Nosenzo, D. (2020). Observability, Social Proximity, and the Erosion of Norm Compliance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3576289

Bicchieri, C., Dimant, E., Gelfand, M., & Sonderegger, S. (2023). Social norms and behavior change: The interdisciplinary research frontier. *Journal of Economic Behavior & Organization*, *205*, A4–A7.

Bicchieri, C., Dimant, E., & Sonderegger, S. (2023). It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion. *Games and Economic Behavior*, *138*, 321–354. https://doi.org/10.1016/j.geb.2023.01.005

Bicchieri, C., Dimant, E., & Xiao, E. (2021). Deviant or wrong? The effects of norm information on the efficacy of punishment. *Journal of Economic Behavior & Organization*, *188*, 209–235.

Bicchieri, C., & Mercier, H. (2014). Norms and Beliefs: How Change Occurs. In M. Xenitidou & B. Edmonds (Hrsg.), *The Complexity of Social Norms* (S. 37–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-05308-0_3

Bicchieri, C., Muldoon, R., & Sontuoso, A. (2018). *Social norms*.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making*, *22*(2), 191–208. https://doi.org/10.1002/bdm.621

Biró, K., & Szalmáné Csete, M. (2022). How Can Development Strategies Foster Agri-digitalisation in the Era of Climate Change? *Periodica Polytechnica Social and Management Sciences*, *31*(1), 9–18. https://doi.org/10.3311/PPso.20411

Björkelo, B., Einarsen, S., Nielsen, M. B., & Matthiesen, S. B. (2011). Silence is golden? Characteristics and experiences of self-reported whistleblowers. *European Journal of*

Work and Organizational Psychology, *20*(2), 206–238. https://doi.org/10.1080/13594320903338884

Bjørkelo, B., & Macko, M. (2012). The stigma of reporting wrongdoing at work: When doing right is perceived as wrong. *Polish Psychological Bulletin*, *43*(2), 70–75.

Bogliacino, F., Charris, R., Gómez, C., & Montealegre, F. (2024). Negative economic shocks and the compliance to social norms. *Judgment and Decision Making*, *19*, e9. https://doi.org/10.1017/jdm.2024.1

Bohnsack, R., Bidmon, C. M., & Pinkse, J. (2022). Sustainability in the digital age: Intended and unintended consequences of digital technologies for sustainable development. *Business Strategy and the Environment*, *31*(2), 599–602. https://doi.org/10.1002/bse.2938

Bolton, G. E., Katok, E., & Zwick, R. (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International Journal of Game Theory*, *27*(2), 269–299. https://doi.org/10.1007/s001820050072

Boo, E., Ng, T., & Shankar, P. G. (2021). Effects of Advice on Auditor Whistleblowing Propensity: Do Advice Source and Advisor Reassurance Matter? *Journal of Business Ethics*, *174*(2), 387–402. https://doi.org/10.1007/s10551-020-04615-0

Böttcher, T. P., Empelmann, S., Weking, J., Hein, A., & Krcmar, H. (2023). Digital sustainable business models: Using digital technology to integrate ecological sustainability into the core of business models. *Information Systems Journal*, *34*(3), 736–761. https://doi.org/10.1111/isj.12436

Brañas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology*, *28*(4), 477–486. https://doi.org/10.1016/j.joep.2006.10.001

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics*, *14*, 375–398.

Brenner, B., & Hartl, B. (2021). The perceived relationship between digitalization and ecological, economic, and social sustainability. *Journal of Cleaner Production*, *315*, 128128. https://doi.org/10.1016/j.jclepro.2021.128128

Brink, A. G., Eller, C. K., & Green, K. Y. (2018). The Effects of Corporate Social Responsibility and Wrongdoer Rank on Whistleblowing. *Accounting and the Public Interest*, *18*(1), 104–128. https://doi.org/10.2308/apin-52240

Broccardo, L., Truant, E., & Dana, L.-P. (2023). The interlink between digitalization, sustainability, and performance: An Italian context. *Journal of Business Research*, *158*, 113621. https://doi.org/10.1016/j.jbusres.2022.113621

Burks, S. V., & Krupka, E. L. (2012). A Multimethod Approach to Identifying Norms and Normative Expectations Within a Corporate Hierarchy: Evidence from the Financial Services Industry. *Management Science*, *58*(1), 203–217. https://doi.org/10.1287/mnsc.1110.1478

Bursztyn, L., Egorov, G., & Fiorin, S. (2020). From Extreme to Mainstream: The Erosion of Social Norms. *American Economic Review*, *110*(11), 3522–3548. https://doi.org/10.1257/aer.20171175

Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review*, *110*(10), 2997–3029. https://doi.org/10.1257/aer.20180975

Bursztyn, L., & Yang, D. Y. (2022). Misperceptions about others. *Annual Review of Economics*, *14*, 425–452.

Butler, J. V., Serra, D., & Spagnolo, G. (2020). Motivating whistleblowers. *Management Science*, *66*(2), 605–621.

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C. F., & Thaler, R. H. (1995). Anomalies: Ultimatums, Dictators and Manners. *Journal of Economic Perspectives*, *9*(2), 209–219. https://doi.org/10.1257/jep.9.2.209

Campoamor, L. M. (2019). There's an App for That: Telecom, Children's Rights, and Conflicting Logics of Corporate Social Responsibility. *American Anthropologist*, *121*(3), 667–679. https://doi.org/10.1111/aman.13273

Carl, K. V. (2021). Corporate Digital Responsibility: Evaluating Privacy and Data Security Activities on Company-level. *INFORMATIK*, *Gesellschaft fur Informatik, Bonn*, 757–771.

Carl, K. V. (2022). The status-quo of companies' data privacy and security communication: An ethical evaluation and future paths. *INFORMATIK*, *Gesellschaft fur Informatik, Bonn*, 195–206. https://doi.org/10.18420/INF2022_18

Carl, K. V. (2023). Data privacy and security in the context of corporate digital responsibility: A scoping review. In *Lecture Notes in Informatics (LNI), Proceedings—Series of the Gesellschaft fur Informatik (GI): Bd. P-337* (S. 523–535). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181099896&partnerID=40&md5=c864aa6ad5910d4b3ebeadb70798f71a

Carl, K. V., & Hinz, O. (2024). What we already know about corporate digital responsibility in IS research: A review and conceptualization of potential CDR activities. *Electronic Markets*, *34*(1), 27. https://doi.org/10.1007/s12525-024-00708-0

Carl, K. V., Mihale-Wilson, C., Zibuschka, J., & Hinz, O. (2024). A consumer perspective on Corporate Digital Responsibility: An empirical evaluation of consumer preferences. *Journal of Business Economics*, *94*(7–8), 979–1024. https://doi.org/10.1007/s11573-023-01142-y

Carl, K. V., Zilcher, T. M. C., & Hinz, O. (2022). Corporate Digital Responsibility and the current Corporate Social Responsibility standard: An analysis of applicability. *Open Identity Summit*, *Gesellschaft für Informatik, Bonn*, 75–86.

Carroll, A. B. (2021). Corporate Social Responsibility: Perspectives on the CSR Construct's Development and Future. *Business & Society*, *60*(6), 1258–1278. https://doi.org/10.1177/00076503211001765

Cassematis, P. G., & Wortley, R. (2013). Prediction of Whistleblowing or Non-reporting Observation: The Role of Personal and Situational Factors. *Journal of Business Ethics*, *117*(3), 615–634. https://doi.org/10.1007/s10551-012-1548-3

Castillo, G., Choo, L., & Grimm, V. (2022). *Do different people report the same social norms?*

Chang, W., Yin, S., Yu, M., Teymurova, V., & Balabeyova, N. (2023). Impact of innovation on Corporate Social Responsibility: Evidence from China. *Economic Analysis and Policy*, *78*, 1185–1194. https://doi.org/10.1016/j.eap.2023.04.018

Chang, Y., Wilding, M., & Shin, M. C. (2017). Determinants of Whistleblowing Intention: Evidence from the South Korean Government. *Public Performance and Management Review*, *40*(4), 676–700. https://doi.org/10.1080/15309576.2017.1318761

Charness, G., Dimant, E., Gneezy, U., & Krupka, E. (2025). *Experimental Methods: Eliciting and Measuring Social Norms*. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5194790

Chatzistamoulou, N. (2023). Is digital transformation the Deus ex Machina towards sustainability transition of the European SMEs? *Ecological Economics*, *206*, 107739. https://doi.org/10.1016/j.ecolecon.2023.107739

Chauhan, C., Akram, M. U., Jahnavi Patky, & Chauhan, A. (2023). Mapping pathways for building resilient supply chains: A systematic literature review. *Journal of Cleaner Production*, *425*, 138701. https://doi.org/10.1016/j.jclepro.2023.138701

Chen, C. X., Nichol, J. E., & Zhou, F. H. (2017). The Effect of Incentive Framing and Descriptive Norms on Internal Whistleblowing. *Contemporary Accounting Research*, *34*(4), 1757–1778. https://doi.org/10.1111/1911-3846.12325

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Chen, L., & Chen, Y. (2023). A Metaorganizations Perspective on Digital Innovation and Corporate Social Responsibility: Evidence from China. *Sustainability*, *15*(14), 11031. https://doi.org/10.3390/su151411031

Chen, W. (2023). Digital economy development, corporate social responsibility and low-carbon innovation. *Corporate Social Responsibility and Environmental Management*, *30*(4), 1664–1679. https://doi.org/10.1002/csr.2443

Cheng, C., & Zhang, M. (2023). Conceptualizing Corporate Digital Responsibility: A Digital Technology Development Perspective. *Sustainability*, *15*(3), 2319. https://doi.org/10.3390/su15032319

Ching, N. T., Ghobakhloo, M., Iranmanesh, M., Maroufkhani, P., & Asadi, S. (2022). Industry 4.0 applications for sustainable manufacturing: A systematic literature review and a roadmap to sustainable development. *Journal of Cleaner Production*, *334*, 130133. https://doi.org/10.1016/j.jclepro.2021.130133

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, *55*(1), 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026.

Clausen, S., Braun, L.-M., & Stieglitz, S. (2023). *Towards More Digital Wellbeing in Knowledge Work—A Signaling Theory Perspective*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2023.561

Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, *26*(4), 2051–2068.

Contini, G., Peruzzini, M., Bulgarelli, S., & Bosi, G. (2023). Developing key performance indicators for monitoring sustainability in the ceramic industry: The role of digitalization and industry 4.0 technologies. *Journal of Cleaner Production*, *414*, 137664. https://doi.org/10.1016/j.jclepro.2023.137664

Cooper, T., Siu, J., & Wei, K. (2015). *Corporate digital responsibility: Doing well by doing good*. https://www.criticaleye.com/inspiring/insights-servfile.cfm?id=4431

Covucci, C., Confetto, M. G., Ključnikov, A., & Panait, M. (2024). Unrevealing the nexus between digital sustainability and corporate digital responsibility: A dual-track systematic literature review towards a framework of corporate digital sustainability. *Technology in Society*, *79*, 102743. https://doi.org/10.1016/j.techsoc.2024.102743

Cox, J. C. (2010). Some issues of methods, theories, and experimental designs. *Journal of Economic Behavior & Organization*, *73*(1), 24–28. https://doi.org/10.1016/j.jebo.2009.01.014

Culiberg, B., & Mihelič, K. K. (2017). The Evolution of Whistleblowing Studies: A Critical Review and Research Agenda. *Journal of Business Ethics*, *146*(4), 787–803. https://doi.org/10.1007/s10551-016-3237-0

Curphy, G. J., Gibson, F. W., Macomber, G., Calhoun, C. J., Wilbanks, L. A., & Burger, M. J. (1998). Situational factors affecting peer reporting intentions at the U.S. Air Force Academy: A scenario-based investigation. *Military Psychology*, *10*(1), 27–43. https://doi.org/10.1207/s15327876mp1001_3

Curtis, M. B., Robertson, J. C., Cockrell, R. C., & Fayard, L. D. (2021). Peer Ostracism as a Sanction Against Wrongdoers and Whistleblowers. *Journal of Business Ethics*, *174*(2), 333–354. https://doi.org/10.1007/s10551-020-04596-0

Ćwiklicki, M., & Wojnarowska, M. (2020). Circular Economy and Industry 4.0: One-Way or Two-way Relationships? *Engineering Economics*, *31*(4), 387–397. https://doi.org/10.5755/j01.ee.31.4.24565

D'Adda, G., Drouvelis, M., & Nosenzo, D. (2016). Norm elicitation in within-subject designs: Testing for order effects. *Journal of Behavioral and Experimental Economics*, *62*, 1–7. https://doi.org/10.1016/j.socec.2016.02.003

D'Adda, G., Dufwenberg, M., Passarelli, F., & Tabellini, G. (2020). Social norms with private values: Theory and experiments. *Games and Economic Behavior*, *124*, 288–304. https://doi.org/10.1016/j.geb.2020.08.012

Dahlsrud, A. (2008). How corporate social responsibility is defined: An analysis of 37 definitions. *Corporate Social Responsibility and Environmental Management*, *15*(1), 1–13. https://doi.org/10.1002/csr.132

Daly, H. E. (1995). On Wilfred Beckerman's Critique of Sustainable Development. *Environmental Values*, *4*(1), 49–55. https://doi.org/10.1177/096327199500400103

Daly, H. E. (1996). *Beyond growth: The economics of sustainable development*. Beacon Press.

Daly, H. E. (2007). *Ecological Economics and Sustainable Development, Selected Essays of Herman Daly*. Edward Elgar Publishing. https://doi.org/10.4337/9781847206947

Dannals, J. E., & Miller, D. T. (2017). Social norm perception in groups with outliers. *Journal of Experimental Psychology: General*, *146*(9), 1342–1359. https://doi.org/10.1037/xge0000336

De Groot, J. I. M., Abrahamse, W., & Jones, K. (2013). Persuasive normative messages: The influence of injunctive and personal norms on using free plastic bags. *Sustainability*, *5*(5), 1829–1844.

De Maria, W., & Jan, C. (1997). Ealing Its Own: The Whistleblower ' s Organization in. *Australien Journal of Social Issues*, *32*(1), 37–59.

De Roeck, K., & Maon, F. (2018). Building the Theoretical Puzzle of Employees' Reactions to Corporate Social Responsibility: An Integrative Conceptual Framework and Research Agenda. *Journal of Business Ethics*, *149*(3), 609–625. https://doi.org/10.1007/s10551-016-3081-2

Del Río Castro, G., González Fernández, M. C., & Uruburu Colsa, Á. (2021). Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *Journal of Cleaner Production*, *280*, 122204. https://doi.org/10.1016/j.jclepro.2020.122204

Demir, S., Gunduz, M. A., Kayikci, Y., & Paksoy, T. (2023). Readiness and Maturity of Smart and Sustainable Supply Chains: A Model Proposal. *Engineering Management Journal*, *35*(2), 181–206. https://doi.org/10.1080/10429247.2022.2050129

Denyer, D., & Tranfield, D. (2009). Producing a Systematic Review. In D. A. Buchanan & A. Bryman (Hrsg.), *The Sage handbook of organizational research methods* (S. 671–689). Sage Publications Ltd.

DesJardins, J. R. (2007). *Business, ethics, and the environment: Imagining a sustainable future*. Pearson/Prentice Hall.

Di Maria, E., De Marchi, V., & Galeazzo, A. (2022). Industry 4.0 technologies and circular economy: The mediating role of supply chain integration. *Business Strategy and the Environment*, *31*(2), 619–632. https://doi.org/10.1002/bse.2940

Dimant, E., & Gesche, T. (2023). Nudging enforcers: How norm perceptions and motives for lying shape sanctions. *PNAS Nexus*, *2*(7), 1–14.

Dimant, E., Van Kleef, G. A., & Shalvi, S. (2020). Requiem for a nudge: Framing effects in nudging honesty. *Journal of Economic Behavior & Organization*, *172*, 247–266.

Dixon, O. (2016). Honest Y Without Fear? Whistleblower Anti-Retaliation Protections in Corporate Code of Conduct. *Melbourne University Law Review*, *40*(1), 168–206.

Djakman, C. D., & Siregar, S. V. (2024). The effect of maturity learn element in Enterprise risk management and corporate social responsibility on the level of digital transformation. *Business Strategy & Development*, *7*(1), e346. https://doi.org/10.1002/bsd2.346

Dmytriyev, S. D., Freeman, R. E., & Hörisch, J. (2021). The Relationship between Stakeholder Theory and Corporate Social Responsibility: Differences, Similarities, and Implications

for Social Issues in Management. *Journal of Management Studies*, *58*(6), 1441–1470. https://doi.org/10.1111/joms.12684

Donaldson, T., & Dunfee, T. W. (1999). *Ties that bind: A social contracts approach to business ethics.* Harvard Business School Press.

Dörr, S., & Lautermann, C. (2024a). Beyond direct stakeholders: The extensive scope of Societal Corporate Digital Responsibility (CDR). In *Organizational Dynamics* (Bd. 53, Nummer 2). https://doi.org/10.1016/j.orgdyn.2024.101057

Dörr, S., & Lautermann, C. (2024b). Beyond direct stakeholders: The extensive scope of Societal Corporate Digital Responsibility (CDR). *Organizational Dynamics*, *53*(2), 101057. https://doi.org/10.1016/j.orgdyn.2024.101057

Dozier, J. B., & Miceli, M. P. (1985). Potential Predictors of Whistle-Blowing A Prosocial Behavior Perspective. *Academy of Management Review*, *10*(4), 823–836.

Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? Framing effects in dictator games. *Experimental Economics*, *16*(3), 349–371. https://doi.org/10.1007/s10683-012-9341-9

Dubreuil, B., & Grégoire, J.-F. (2013). Are moral norms distinct from social norms? A critical assessment of Jon Elster and Cristina Bicchieri. *Theory and Decision*, *75*, 137–152.

Dungan, J. A., Young, L., & Waytz, A. (2019). Journal of Experimental Social Psychology The power of moral concerns in predicting whistleblowing decisions. *Journal of Experimental Social Psychology*, *85*(September 2018), 103848. https://doi.org/10.1016/j.jesp.2019.103848

Dwivedi, A., Chowdhury, P., Agrawal, D., Paul, S. K., & Shi, Y. (2023). Antecedents of digital supply chains for a circular economy: A sustainability perspective. *Industrial Management & Data Systems*, *123*(6), 1690–1716. https://doi.org/10.1108/IMDS-05-2022-0273

Dwivedi, A., & Paul, S. K. (2022). A framework for digital supply chains in the era of circular economy: Implications on environmental sustainability. *Business Strategy and the Environment*, *31*(4), 1249–1274. https://doi.org/10.1002/bse.2953

Dworkin, T. M., & Baucus, M. S. (1998). Internal vs. External Whistleblowers: A Comparison of Whistleblowering Processes. *Journal of Business Ethics*, *17*, 1281–1298.

Dyck, A., Morse, A., & Zingales, L. (2010). Who blows the whistle on corporate fraud? *The Journal of Finance*, *65*(6), 2213–2253.

Elkington, J. (1998). ACCOUNTING FOR THE TRIPLE BOTTOM LINE. *MESURING BUSINESS EXCELLENCE*, *2*(3), 18–22. https://doi.org/10.1108/eb025539

Elliott, K., & Copilah-Ali, J. (2024a). Implementing corporate digital responsibility (CDR): Tackling wicked problems for the digital era: Pilot study insights. In *Organizational Dynamics* (Bd. 53, Nummer 2). https://doi.org/10.1016/j.orgdyn.2024.101040

Elliott, K., & Copilah-Ali, J. (2024b). Implementing corporate digital responsibility (CDR): Tackling wicked problems for the digital era: Pilot study insights. *Organizational Dynamics*, *53*(2), 101040. https://doi.org/10.1016/j.orgdyn.2024.101040

Elliott, K., Price, R., Shaw, P., Spiliotopoulos, T., Ng, M., Coopamootoo, K., & Van Moorsel, A. (2021). Towards an Equitable Digital Society: Artificial Intelligence (AI) and Corporate Digital Responsibility (CDR). *Society*, *58*(3), 179–188. https://doi.org/10.1007/s12115-021-00594-8

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, *18*(3), 199–212. https://doi.org/10.1002/bdm.495

Epp, L., Foit, D., Kürpick, C., Plaß, S., Scholz, T., & Schreiner, N. (2024). Developing A Conceptual Model for Strategic Integration of Sustainability and Digitalization in Manufacturing. *2024 IEEE International Conference on Technology, Informatics, Management, Engineering and Environment (TIME-E)*, 100–105. https://doi.org/10.1109/TIME-E62724.2024.10919721

Esposito, P., & Ricci, P. (2021). Cultural organizations, digital Corporate Social Responsibility and stakeholder engagement in virtual museums: A multiple case study. How digitization is influencing the attitude toward CSR. *Corporate Social Responsibility and Environmental Management*, *28*(2), 953–964. https://doi.org/10.1002/csr.2074

Etter, M., Fieseler, C., & Whelan, G. (2019). Sharing Economy, Sharing Responsibility? Corporate Social Responsibility in the Digital Age. *Journal of Business Ethics*, *159*(4), 935–942. https://doi.org/10.1007/s10551-019-04212-w

Falkenberg, L., & Herremans, I. (1995). Ethical behaviours in organizations: Directed by the formal or informal systems? *Journal of Business Ethics*, *14*(2), 133–143. https://doi.org/10.1007/BF00872018

Fallucchi, F., & Nosenzo, D. (2022). The coordinating power of social norms. *Experimental Economics*, *25*(1), 1–25. https://doi.org/10.1007/s10683-021-09717-8

Famularo, J. (2023). Corporate social responsibility communication in the ICT sector: Digital issues, greenwashing, and materiality. *International Journal of Corporate Social Responsibility*, *8*(1), 8. https://doi.org/10.1186/s40991-023-00082-8

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, *25*(2), 63–87.

Fehr, E., & Schmidt, K. M. (2006). Chapter 8 The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories. In *Handbook of the Economics of Giving, Altruism and Reciprocity* (Bd. 1, S. 615–691). Elsevier. https://doi.org/10.1016/S1574-0714(06)01008-6

Fellner, G., Sausgruber, R., & Traxler, C. (2013). Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information. *Journal of the European Economic Association*, *11*(3), 634–660. https://doi.org/10.1111/jeea.12013

Feroz, A. K., Zo, H., Eom, J., & Chiravuri, A. (2023). Identifying organizations' dynamic capabilities for sustainable digital transformation: A mixed methods study. *Technology in Society*, *73*, 1–13. https://doi.org/10.1016/j.techsoc.2023.102257

Ferrante, M., Vitti, M., Facchini, F., & Sassanelli, C. (2024). Mapping the relations between the circular economy rebound effects dimensions: A systematic literature review. *Journal of Cleaner Production*, *456*, 142399. https://doi.org/10.1016/j.jclepro.2024.142399

Ferreira, J. J., Lopes, J. M., Gomes, S., & Rammal, H. G. (2023). Industry 4.0 implementation: Environmental and social sustainability in manufacturing multinational enterprises. *Journal of Cleaner Production*, *404*, 136841. https://doi.org/10.1016/j.jclepro.2023.136841

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Findik, D., Tirgil, A., & Özbuğday, F. C. (2023). Industry 4.0 as an enabler of circular economy practices: Evidence from European SMEs. *Journal of Cleaner Production*, *410*, 137281. https://doi.org/10.1016/j.jclepro.2023.137281

Flyverbom, M., Deibert, R., & Matten, D. (2019). The Governance of Digital Technology, Big Data, and the Internet: New Roles and Responsibilities for Business. *Business & Society*, *58*(1), 3–19. https://doi.org/10.1177/0007650317727540

Freeman, R. E. (2010). *Strategic management: A stakeholder approach.* Cambridge university press.

Freeman, R. E., Dmytriyev, S., & Strand, R. G. (2017). Managing for stakeholders in the digital age. In *Corporate Social Responsibility: Strategy, Communication, Governanc* (S. 136–153). Cambridge University Press.

Freeman, R. E., & Elms, H. (2023). The social responsibility of business is to create value for stakeholders. In *R. Edward Freeman's Selected Works on Stakeholder Theory and Business Ethics* (S. 807–810). Cham: Springer International Publishing.

Fromell, H., Nosenzo, D., Owens, T., & Tufano, F. (2021). One size does not fit all: Plurality of social norms and saving behavior in Kenya. *Journal of Economic Behavior & Organization*, *192*, 73–91. https://doi.org/10.1016/j.jebo.2021.09.028

Fu, S., Yu, Y., Su, I.-H., Ling, Z., Tan, K. H., & Ma, R. (2023). The influence of corporate social responsibility information transparency on the consumption of green agricultural products on digital platforms. *International Journal of Logistics Research and Applications*, 1–18. https://doi.org/10.1080/13675567.2023.2242284

Gächter, S., Gerhards, L., & Nosenzo, D. (2017). The importance of peers for compliance with norms of fair sharing. *European Economic Review*, *97*, 72–86. https://doi.org/10.1016/j.euroecorev.2017.06.001

Gächter, S., & Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, *13*(3), 364–377. https://doi.org/10.1007/s10683-010-9246-4

Gagnon, M., & Perron, A. (2020). Whistleblowing: A concept analysis. *Nursing & Health Sciences*, *22*(2), 381–389.

Gao, J., Greenberg, R., & Wong-On-Wing, B. (2015). Whistleblowing intentions of lower-level employees: The effect of reporting channel, bystanders, and wrongdoer power status. *Journal of Business Ethics*, *126*(1), 85–99. https://doi.org/10.1007/s10551-013-2008-4

Gao, L., & Brink, A. G. (2017). Whistleblowing studies in accounting research: A review of experimental studies on the determinants of whistleblowing. *Journal of Accounting Literature*, *38*(May), 1–13. https://doi.org/10.1016/j.acclit.2017.05.001

Garriga, E., & Mele, D. (2004). Corporate Social Responsibility Theories: Mapping the Territory. *Journal of Business Ethics*, *53*, 51–71.

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, *57*(2), 321–326.

George, G., Merrill, R. K., & Schillebeeckx, S. J. D. (2021). Digital Sustainability and Entrepreneurship: How Digital Innovations Are Helping Tackle Climate Change and Sustainable Development. *Entrepreneurship Theory and Practice*, *45*(5), 999–1027. https://doi.org/10.1177/1042258719899425

Ghobakhloo, M., Asadi, S., Iranmanesh, M., Foroughi, B., Mubarak, M. F., & Yadegaridehkordi, E. (2023a). Intelligent automation implementation and corporate sustainability performance: The enabling role of corporate social responsibility strategy. *Technology in Society*, *74*, 102301. https://doi.org/10.1016/j.techsoc.2023.102301

Ghobakhloo, M., Asadi, S., Iranmanesh, M., Foroughi, B., Mubarak, M. F., & Yadegaridehkordi, E. (2023b). Intelligent automation implementation and corporate sustainability performance: The enabling role of corporate social responsibility strategy. *Technology in Society*, *74*, 102301. https://doi.org/10.1016/j.techsoc.2023.102301

Ghobakhloo, M., Iranmanesh, M., Grybauskas, A., Vilkas, M., & Petraitė, M. (2021). Industry 4.0, innovation, and sustainable development: A systematic review and a roadmap to sustainable innovation. *Business Strategy and the Environment*, *30*(8), 4237–4257. https://doi.org/10.1002/bse.2867

Ghobakhloo, M., Iranmanesh, M., Morales, M. E., Nilashi, M., & Amran, A. (2023). Actions and approaches for enabling Industry 5.0-driven sustainable industrial transformation: A strategy roadmap. *Corporate Social Responsibility and Environmental Management*, *30*(3), 1473–1494. https://doi.org/10.1002/csr.2431

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, *76*(1), 129–142. https://doi.org/10.1037/0022-3514.76.1.129

Giddings, B., Hopwood, B., & O'brien, G. (2002). Environment, economy and society: Fitting them together into sustainable development. *Sustainable development*, *10*(4), 187–196.

Gilbert, D. U., Schrage, S., & Behnam, M. (2024). Advancing the Moral legitimacy of digital platforms as gatekeepers: A critical analysis from a political corporate social responsibility perspective. *Journal of Business Economics*, *94*(7–8), 1115–1145. https://doi.org/10.1007/s11573-024-01200-z

Girrbach, P. (2021). Corporate Responsibility in the Context of Digitalization. *Tehnički Glasnik*, *15*(3), 422–428. https://doi.org/10.31803/tg-20210710142357

Glazer, M. P., & Glazer, A. P. (1989). *The Whistleblowers: Exposing Corruption in Government and Industry*. Basic Books.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, *25*(4), 191–210. https://doi.org/10.1257/jep.25.4.191

Gneezy, U., Saccardo, S., & Van Veldhuizen, R. (2019). Bribery: Behavioral Drivers of Distorted Decisions. *Journal of the European Economic Association*, *17*(3), 917–946. https://doi.org/10.1093/jeea/jvy043

Goddiksen, M. P., Quinn, U., Kovács, N., Lund, T. B., Sandøe, P., Varga, O., & Willum Johansen, M. (2021). Good friend or good student? An interview study of perceived conflicts between personal and academic integrity among students in three European countries. *Accountability in Research*, *28*(4), 247–264. https://doi.org/10.1080/08989621.2020.1826319

Goede, R. (2021). Sustainable business intelligence systems: Modelling for the future. *Systems Research and Behavioral Science*, *38*(5), 685–695. https://doi.org/10.1002/sres.2813

Goldsmith, D. J. (2004). *Communicating social support*. Cambridge University Press.

González-Ramírez, R., Gasco, J., & Llopis, J. (2024). Digitalisation and sustainability: Their role in corporate social responsibility through innovation. *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-04-2024-0410

Görges, L., & Nosenzo, D. (2020). Measuring Social Norms in Economics: Why It Is Important and How It Is Done. *Analyse & Kritik*, *42*(2), 285–312. https://doi.org/10.1515/auk-2020-0012

Govindan, K. (2024). Theory Building Through Corporate Social Responsibility 4.0 for Achieving SDGs: A Practical Step Toward Integration of Digitalization With Practice-Based View and Social Good Theory. *IEEE Transactions on Engineering Management*, *71*, 2103–2120. https://doi.org/10.1109/TEM.2022.3155247

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Greenberger, D. B., Miceli, M. P., & Cohen, D. J. (1987). Oppositionists and group norms: The reciprocal influence of whistle-blowers and co-workers. *Journal of Business Ethics*, *6*(7), 527–542. https://doi.org/10.1007/BF00383744

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.

Grunwald, G. (2022). Sustainability co-creation in digitalized global value chains. *Strategic Change*, *31*(1), 19–29. https://doi.org/10.1002/jsc.2477

Guandalini, I. (2022). Sustainability through digital transformation: A systematic literature review for research guidance. *Journal of Business Research*, *148*, 456–471. https://doi.org/10.1016/j.jbusres.2022.05.003

Gundlach, M. J., Douglas, S. C. ., & Martinko, M. J. . (2003). The Decision to Blow the Whistle: A Social Information Processing Framework. *Academy of Management Review*, *28*(1), 107–123.

Guo, Y., Zou, H., Liu, Z., & Liu, B. (2024). Research on differential game of platform corporate social responsibility governance strategy considering user and public scrutiny. *PLOS ONE*, *19*(4), e0301632. https://doi.org/10.1371/journal.pone.0301632

Gupta, A., & Singh, R. K. (2021). Applications of emerging technologies in logistics sector for achieving circular economy goals during COVID 19 pandemic: Analysis of critical success factors. *International Journal of Logistics Research and Applications*, 1–22. https://doi.org/10.1080/13675567.2021.1985095

Haftor, D. M., & Climent, R. C. (2021). CO2 reduction through digital transformation in long-haul transportation: Institutional entrepreneurship to unlock product-service system innovation. *Industrial Marketing Management*, *94*, 115–127. https://doi.org/10.1016/j.indmarman.2020.08.022

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, 14–31.

Hartley, N., Kunz, W., & Tarbit, J. (2024). The corporate digital responsibility (CDR) calculus: How and why organizations reconcile digital and ethical trade-offs for growth. *Organizational Dynamics*, *53*(2), 101056. https://doi.org/10.1016/j.orgdyn.2024.101056

Hashim, M. A. M., & et al. (2022). Emergent Strategy in Higher Education: Postmodern Digital and the Future? *Administrative Sciences*.

Hassink, H., De Vries, M., & Bollen, L. (2007). A content analysis of whistleblowing policies of leading European companies. *Journal of Business Ethics*, *75*(1), 25–44. https://doi.org/10.1007/s10551-006-9236-9

Hauser, O. P., Linos, E., & Rogers, T. (2017). Innovation with field experiments: Studying organizational behaviors in actual organizations. *Research in Organizational Behavior*, *37*, 185–198. https://doi.org/10.1016/j.riob.2017.10.004

He, Z., Kuai, L., & Wang, J. (2023). Driving mechanism model of enterprise green strategy evolution under digital technology empowerment: A case study based on Zhejiang Enterprises. *Business Strategy and the Environment*, *32*(1), 408–429. https://doi.org/10.1002/bse.3138

Heal, G. (2005). Corporate Social Responsibility: An Economic and Financial Framework. *The Geneva Papers on Risk and Insurance - Issues and Practice*, *30*(3), 387–409. https://doi.org/10.1057/palgrave.gpp.2510037

Heinrich, T., & Weimann, J. (2013). A note on reciprocity and modified dictator games. *Economics Letters*, *121*(2), 202–205. https://doi.org/10.1016/j.econlet.2013.08.004

Herden, C. J., Alliu, E., Cakici, A., Cormier, T., Deguelle, C., Gambhir, S., Griffiths, C., Gupta, S., Kamani, S. R., Kiratli, Y.-S., Kispataki, M., Lange, G., Moles De Matos, L., Tripero Moreno, L., Betancourt Nunez, H. A., Pilla, V., Raj, B., Roe, J., Skoda, M., … Edinger-Schons, L. M. (2021). "Corporate Digital Responsibility": New corporate responsibilities in the digital age. *Sustainability Management Forum | NachhaltigkeitsManagementForum*, *29*(1), 13–29. https://doi.org/10.1007/s00550-020-00509-x

Herschel, R., & Miori, V. M. (2017). Ethics & Big Data. *Technology in Society*, *49*, 31–36. https://doi.org/10.1016/j.techsoc.2017.03.003

Hina, M., Chauhan, C., Kaur, P., Kraus, S., & Dhir, A. (2022). Drivers and barriers of circular economy business models: Where we are now, and where we are heading. *Journal of Cleaner Production*, *333*, 130049. https://doi.org/10.1016/j.jclepro.2021.130049

Hollinger, R. C., & Clark, J. P. (1982). Formal and Informal Social Controls of Employee Deviance Linked references are available on JSTOR for this article: Formal and Informal Social Controls of Employee Deviance *. *Sociological Quarterly*, *23*(3), 333–343.

Horne, C., & Mollborn, S. (2020). Norms: An Integrated Framework. *Annual Review of Sociology*, *46*(1), 467–487. https://doi.org/10.1146/annurev-soc-121919-054658

Hristov, I., & Appolloni, A. (2022). Stakeholders' engagement in the business strategy as a key driver to increase companies' performance: Evidence from managerial and stakeholders' practices. *Business Strategy and the Environment*, *31*(4), 1488–1503. https://doi.org/10.1002/bse.2965

Hu, Y., & Liu, Q. (2023). Local Digital Economy and Corporate Social Responsibility. *Sustainability*, *15*(11), 8487. https://doi.org/10.3390/su15118487

Huang, G., & Shen, L. (2024). The Bidirectional Relationship between Digital Transformation and Corporate Social Responsibility: A Legitimacy Perspective. *Sustainability*, *16*(7), 3029. https://doi.org/10.3390/su16073029

Huang, R., & Wei, J. (2023). Does CEOs' green experience affect environmental corporate social responsibility? Evidence from China. *Economic Analysis and Policy*, *79*, 205–231. https://doi.org/10.1016/j.eap.2023.06.012

Iriberri, N., & Rey-Biel, P. (2011). The role of role uncertainty in modified dictator games. *Experimental Economics*, *14*(2), 160–180. https://doi.org/10.1007/s10683-010-9261-5

Isensee, C., Teuteberg, F., Griese, K.-M., & Topi, C. (2020). The relationship between organizational culture, sustainability, and digitalization in SMEs: A systematic review. *Journal of Cleaner Production*, *275*, 1–19. https://doi.org/10.1016/j.jclepro.2020.122944

Islam, A., Wahab, S. A., & Latiff, A. S. A. (2022). Annexing a smart sustainable business growth model for small and medium enterprises (SMEs). *World Journal of Entrepreneurship, Management and Sustainable Development*, *18*(2), 22–46. https://doi.org/10.47556/J.WJEMSD.18.2.2022.2

Iwai, T., Yeung, L., & Artes, R. (2019). Voice or silence: Antecedents of whistleblowing intentions. *RAUSP Management Journal*, *56*(2), 186–201. https://doi.org/10.1108/RAUSP-06-2020-0126

Iwai, T., Yeung, L., & Artes, R. (2021). Voice or silence: Antecedents of whistleblowing intentions. *RAUSP Management Journal*, *56*(2), 186–201. https://doi.org/10.1108/RAUSP-06-2020-0126

Jackson, D., Peters, K., Andrew, S., Edenborough, M., Luck, L., Salamonson, Y., Weaver, R., & Wilkes, L. (2010). Trial and retribution: A qualitative study of whistleblowing and workplace relationships in nursing. *Contemporary Nurse*, *36*(1–2), 34–44. https://doi.org/10.5172/conu.2010.36.1-2.034

Jamwal, A., Agrawal, R., & Sharma, M. (2022). Deep learning for manufacturing sustainability: Models, applications in Industry 4.0 and implications. *International Journal of*

Information Management Data Insights, *2*(2), 1–13. https://doi.org/10.1016/j.jjimei.2022.100107

Jelovac, D., Ljubojević, Č., & Ljubojević, L. (2022). HPC in business: The impact of corporate digital responsibility on building digital trust and responsible corporate digital governance. *Digital Policy, Regulation and Governance*, *24*(6), 485–497. https://doi.org/10.1108/DPRG-11-2020-0164

Jiang, W., Wu, J., & Yang, X. (2023). Does digitization drive corporate social responsibility? *International Review of Economics & Finance*, *88*, 14–26. https://doi.org/10.1016/j.iref.2023.06.010

Jones, P., & Comfort, D. (2021). Corporate Digital Responsibility Challenges for Sports Betting Companies. *Journal of Gambling Issues*, *48*, 202–211. https://doi.org/10.4309/jgi.2021.48.9

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review*, *16*(2), 366–395.

Jung, H., Bae, J., & Kim, H. (2022). The effect of corporate social responsibility and corporate social irresponsibility: Why company size matters based on consumers' need for self-expression. *Journal of Business Research*, *146*, 146–154. https://doi.org/10.1016/j.jbusres.2022.03.024

Jurgens, M., Berthon, P., Edelman, L., & Pitt, L. (2016). Social media revolutions: The influence of secondary stakeholders. *Business Horizons*, *59*(2), 129–136. https://doi.org/10.1016/j.bushor.2015.11.010

Kaplan, S. E., Pope, K. R., & Samuels, J. A. (2010). The effect of social confrontation on individuals' intentions to internally report fraud. *Behavioral Research in Accounting*, *22*(2), 51–67. https://doi.org/10.2308/bria.2010.22.2.51

Kaptein, M. (2011). Toward Effective Codes: Testing the Relationship with Unethical Behavior. *Journal of Business Ethics*, *99*(2), 233–251. https://doi.org/10.1007/s10551-010-0652-5

Kärpänen, T. (2022, Juli 20). Corporate Digital Responsibility and Accessibility in Digital Services. *Proceedings of the 16th International Conference on Interfaces and Human Computer Interaction 2022 and 15th International Conference on Game and Entertainment Technologies 2022*. https://doi.org/10.33965/IHCI_GET2022_202205L012

Kassas, B., & Palma, M. A. (2019). Self-serving biases in social norm compliance. *Journal of Economic Behavior & Organization*, *159*, 388–408. https://doi.org/10.1016/j.jebo.2019.02.010

Keenan, J. P. (2000). *Blowing the Whistle on Less Serious Forms of Fraud: A Study of Executives and Managers 1*. *12*(4).

Keenan, J. P. (2007). Comparing Chinese and American Managers on Whistleblowing. *Employee Responsibilities and Rights Journal*, *19*(2), 85–94. https://doi.org/10.1007/s10672-007-9036-0

Khan, S. A. R., Sheikh, A. A., & Tahir, M. S. (2024). Corporate social responsibility—an antidote for sustainable business performance: Interconnecting role of digital technologies, employee eco-behavior, and tax avoidance. *Environmental Science and Pollution Research*, *31*(3), 4365–4383. https://doi.org/10.1007/s11356-023-31377-9

Khan, S. R., & Howe, L. C. (2021). Concern for the Transgressor's Consequences: An Explanation for Why Wrongdoing Remains Unreported. *Journal of Business Ethics*, *173*(2), 325–344. https://doi.org/10.1007/s10551-020-04568-4

Khattak, A., & Yousaf, Z. (2022). Digital Social Responsibility towards Corporate Social Responsibility and Strategic Performance of Hi-Tech SMEs: Customer Engagement as a Mediator. *Sustainability*, *14*(1), 131. https://doi.org/10.3390/su14010131

Kimbrough, E. O., & Vostroknutov, A. (2016). Norms Make Preferences Social. *Journal of the European Economic Association*, *14*(3), 608–638. https://doi.org/10.1111/jeea.12152

Kimbrough, E. O., & Vostroknutov, A. (2018). A portable method of eliciting respect for social norms. *Economics Letters*, *168*, 147–150. https://doi.org/10.1016/j.econlet.2018.04.030

Kluiters, L., Srivastava, M., & Tyll, L. (2023a). The impact of digital trust on firm value and governance: An empirical investigation of US firms. In *Society and Business Review* (Bd. 18, Nummer 1, S. 71–103). https://doi.org/10.1108/SBR-07-2021-0119

Kluiters, L., Srivastava, M., & Tyll, L. (2023b). The impact of digital trust on firm value and governance: An empirical investigation of US firms. *Society and Business Review*, *18*(1), 71–103. https://doi.org/10.1108/SBR-07-2021-0119

Knoll, M., & Van Dick, R. (2013). Do I Hear the Whistle…? A First Attempt to Measure Four Forms of Employee Silence and Their Correlates. *Journal of Business Ethics*, *113*(2), 349–362. https://doi.org/10.1007/s10551-012-1308-4

Knopf, T., & Pick, D. (2023). Corporate Responsibility for Digital Innovation: A systematic Review of the Literature. *European Conference on Innovation and Entrepreneurship*, *18*(1), 469–477. https://doi.org/10.34190/ecie.18.1.1601

Kohlberg, L. (1964). Development of Moral Character and Moral Ideology. In M. L. Hoffman & L. W. Hoffman (Hrsg.), *Review of Child Development Research* (S. 383–431). Russell Sage Foundation.

Kolyperas, D., Anagnostopoulos, C., Pavlopoulou, I., Manoli, A. E., & Chadwick, S. (2024). Corporate social and digital responsibility in esports. *Internet Research*. https://doi.org/10.1108/INTR-01-2024-0134

Kong, D., & Liu, B. (2023). Digital Technology and Corporate Social Responsibility: Evidence from China. *Emerging Markets Finance and Trade*, *59*(9), 2967–2993. https://doi.org/10.1080/1540496X.2023.2199122

König-Kersting, C. (2024). On the robustness of social norm elicitation. *Journal of the Economic Science Association*. https://doi.org/10.1007/s40881-024-00178-2

Koutras, N. (2019). The Evolving Role of Commercial Publishers and the Future of Open Access Repositories: The Potential of Corporate Social Responsibility. *Publishing Research Quarterly*, *35*(3), 391–417. https://doi.org/10.1007/s12109-019-09644-w

Kristoffersen, E., Blomsma, F., Mikalef, P., & Li, J. (2020). The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies. *Journal of Business Research*, *120*, 241–261. https://doi.org/10.1016/j.jbusres.2020.07.044

Kristoffersen, E., Mikalef, P., Blomsma, F., & Li, J. (2021). The effects of business analytics capability on circular economy implementation, resource orchestration capability, and

firm performance. *International Journal of Production Economics*, *239*, 108205. https://doi.org/10.1016/j.ijpe.2021.108205

Krlev, G., Hannigan, T., & Spicer, A. (2025). What Makes a Good Review Article? Empirical Evidence From Management and Organization Research. *Academy of Management Annals*, *19*(1), 376–403. https://doi.org/10.5465/annals.2021.0051

Krupka, E. L., & Weber, R. A. (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, *30*(3), 307–320.

Krupka, E. L., & Weber, R. A. (2013). Identifying Social Norms Using Coordination Games: Why does Dictator Game Sharing Vary? *Journal of the European Economic Association*, *11*(3), 495–524. https://doi.org/10.1111/jeea.12006

Kuang, J., & Bicchieri, C. (2024a). *How do people interpret social norm messages? Semantic ambiguity and pragmatic inferences*. https://doi.org/10.31234/osf.io/zmgjn

Kuang, J., & Bicchieri, C. (2024b). Language matters: How normative expressions shape norm perception and affect norm compliance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1897), 20230037. https://doi.org/10.1098/rstb.2023.0037

Kujala, J., Sachs, S., Leinonen, H., Heikkinen, A., & Laude, D. (2022). Stakeholder Engagement: Past, Present, and Future. *Business & Society*, *61*(5), 1136–1196. https://doi.org/10.1177/00076503211066595

Kumar, N., Kumar, G., & Singh, R. K. (2021). Big data analytics application for sustainable manufacturing operations: Analysis of strategic factors. *Clean Technologies and Environmental Policy*, *23*(3), 965–989. https://doi.org/10.1007/s10098-020-02008-5

Kumar, V., Vrat, P., & Shankar, R. (2022). Factors Influencing the Implementation of Industry 4.0 for Sustainability in Manufacturing. *Global Journal of Flexible Systems Management*, *23*(4), 453–478. https://doi.org/10.1007/s40171-022-00312-1

Kunicke, K. (2018). *Gesellschaftliche Verantwortung auch für Bits und Bytes. Telekom*. https://www.telekom.com/de/verantwortung/news/gesellschaftliche-verantwortung-auch-fuer-bits-und-bytes-544792

Kunz, W. H., & Wirtz, J. (2024). Corporate digital responsibility (CDR) in the age of AI: Implications for interactive marketing. *Journal of Research in Interactive Marketing*, *18*(1), 31–37. https://doi.org/10.1108/JRIM-06-2023-0176

Kürpick, C., Schreiner, N., Krauß-Kodytek, L., Kühn, A., Plass, S., & Scholz, T. (2024). Capabilities for the Strategic Alignment of Sustainability and Digitalization in Manufacturing: Insights from Theory and Practice. *2024 IEEE 65th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, 1–6. https://doi.org/10.1109/ITMS64072.2024.10741924

Lane, T., Nosenzo, D., & Sonderegger, S. (2023). Law and Norms: Empirical Evidence. *American Economic Review*, *113*(5), 1255–1293. https://doi.org/10.1257/aer.20210970

Latan, H., Ringle, C. M., & Jabbour, C. J. C. (2018). Whistleblowing intentions among public accountants in indonesia: Testing for the moderation effects. *Journal of Business Ethics*, *152*(2), 573–588. https://doi.org/10.1007/s10551-016-3318-0

Latapí Agudelo, M. A., Jóhannsdóttir, L., & Davídsdóttir, B. (2019). A literature review of the history and evolution of corporate social responsibility. *International Journal of Corporate Social Responsibility*, *4*(1), 1. https://doi.org/10.1186/s40991-018-0039-y

Lee, G., & Fargher. (2013). Companies' use of whistle-blowing to detect fraud: An examination of corporate whistle-blowing policies. *Journal of Business Ethics*, *114*, 283–295.

Lee, G., & Xiao, X. (2018). Whistleblowing on accounting-related misconduct: A synthesis of the literature. *Journal of Accounting Literature*, *41*, 22–46.

Legner, C., Eymann, T., Hess, T., Matt, C., Böhmann, T., Drews, P., Mädche, A., Urbach, N., & Ahlemann, F. (2017a). Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. *Business & Information Systems Engineering*, *59*(4), 301–308. https://doi.org/DOI 10.1007/s12599-017-0484-2

Legner, C., Eymann, T., Hess, T., Matt, C., Böhmann, T., Drews, P., Mädche, A., Urbach, N., & Ahlemann, F. (2017b). Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. *Business & Information Systems Engineering*, *59*(4), 301–308. https://doi.org/10.1007/s12599-017-0484-2

Lerman, L. V., Benitez, G. B., Müller, J. M., De Sousa, P. R., & Frank, A. G. (2022). Smart green supply chain management: A configurational approach to enhance green performance through digital transformation. *Supply Chain Management: An International Journal*, *27*(7), 147–176. https://doi.org/10.1108/SCM-02-2022-0059

Lewis, D. (2022). Stigma and Whistleblowing: Should Punitive Damages be Available in Retaliation Cases ? *Industrial Law Journal*, *51*(1), 62–83. https://doi.org/10.1093/indlaw/dwaa032

Li, B., Guo, F., Xu, L., & Meng, S. (2024). Fintech business and corporate social responsibility practices. *Emerging Markets Review*, *59*, 101105. https://doi.org/10.1016/j.ememar.2023.101105

Li, H., Lu, L., Lin, Z., & Meng, T. (2024). Digital innovation and corporate social responsibility performance: Evidence from firms' digital patents. *Technological Forecasting and Social Change*, *207*, 123626. https://doi.org/10.1016/j.techfore.2024.123626

Li, L. (2022). Digital transformation and sustainable performance: The moderating role of market turbulence. *Industrial Marketing Management*, *104*, 28–37. https://doi.org/10.1016/j.indmarman.2022.04.007

Li, M. (2021). The Synergistic Effects of Solutions Journalism and Corporate Social Responsibility Advertising. *Digital Journalism*, *9*(3), 336–363. https://doi.org/10.1080/21670811.2020.1840407

Li, X., Wu, T., Zhang, H., & Yang, D. (2022). Digital Technology Adoption and Sustainable Development Performance of Strategic Emerging Industries: The Mediating Role of Digital Technology Capability and the Moderating Role of Digital Strategy. *Journal of Organizational and End User Computing*, *34*(8), 1–18. https://doi.org/10.4018/JOEUC.315645

Liao, M., Gao, Z., Zhou, J., & Li, D. (2023). Business model innovation driven by corporate social responsibility in a digital innovation ecosystem: Evidence from Chinese manufacturers. *Managerial and Decision Economics*, mde.3998. https://doi.org/10.1002/mde.3998

Lin, J., Zheng, Q., & Benitez, J. (2024). Impact of digital orientation and corporate social responsibility activities on the alliance relationship stability in contract farming. *Information & Management*, *61*(4), 103958. https://doi.org/10.1016/j.im.2024.103958

Lin, Y.-T., Liu, N.-C., & Lin, J.-W. (2022). Firms' adoption of CSR initiatives and employees' organizational commitment: Organizational CSR climate and employees' CSR-induced attributions as mediators. *Journal of Business Research*, *140*, 626–637. https://doi.org/10.1016/j.jbusres.2021.11.028

Lindman, J., Makinen, J., & Kasanen, E. (2023). Big Tech's power, political corporate social responsibility and regulation. *Journal of Information Technology*, *38*(2), 144–159. https://doi.org/10.1177/02683962221113596

Lin-Hi, N., & Müller, K. (2013). The CSR bottom line: Preventing corporate social irresponsibility. *Journal of Business Research*, *66*(10), 1928–1936. https://doi.org/10.1016/j.jbusres.2013.02.015

Liu, H., Han, P., & Wang, S. (2024). Enhancing corporate social responsibility in the digital economy era: Evidence from China. *Heliyon*, *10*(1), e23459. https://doi.org/10.1016/j.heliyon.2023.e23459

Liu, Q., Trevisan, A. H., Yang, M., & Mascarenhas, J. (2022). A framework of digital technologies for the circular economy: Digital functions and mechanisms. *Business Strategy and the Environment*, *31*(5), 2171–2192. https://doi.org/10.1002/bse.3015

Liu, W., Liu, T., Tang, O., Lee, P. T. W., & Chen, Z. (2024). The impact of digital supply chain announcements disclosing corporate social responsibility information on stock market value. *Industrial Management & Data Systems*, *124*(2), 724–760. https://doi.org/10.1108/IMDS-03-2023-0189

Liyanaarachchi, G., Deshpande, S., & Weaven, S. (2021). Market-oriented corporate digital responsibility to manage data vulnerability in online banking. *International Journal of Bank Marketing*, *39*(4), 571–591. https://doi.org/10.1108/IJBM-06-2020-0313

Liyanarachchi, G., & Newdick, C. (2009a). The impact of moral reasoning and retaliation on whistle-blowing: New Zealand evidence. *Journal of Business Ethics*, *89*(1), 37–57. https://doi.org/10.1007/s10551-008-9983-x

Liyanarachchi, G., & Newdick, C. (2009b). The impact of moral reasoning and retaliation on whistle-blowing: New Zealand evidence. *Journal of Business Ethics*, *89*, 37–57.

Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021a). Corporate digital responsibility. In *Journal of Business Research* (Bd. 122, S. 875–888). https://doi.org/10.1016/j.jbusres.2019.10.006

Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021b). Corporate digital responsibility. *Journal of Business Research*, *122*, 875–888. https://doi.org/10.1016/j.jbusres.2019.10.006

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Lopes de Sousa Jabbour, A. B., Jabbour, C. J. C., Godinho Filho, M., & Roubaud, D. (2018). Industry 4.0 and the circular economy: A proposed research agenda and original roadmap

for sustainable operations. *Annals of Operations Research*, *270*(1–2), 273–286. https://doi.org/10.1007/s10479-018-2772-8

López Jiménez, D., Dittmar, E. C., & Vargas Portillo, J. P. (2021). New Directions in Corporate Social Responsibility and Ethics: Codes of Conduct in the Digital Environment. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-021-04753-z

López-Nicolás, C., Meroño-Cerdán, Á. L., Heikkilä, M., & Bouwman, H. (2024). Untangling business model innovation in family firms: Socioemotional wealth and corporate social responsibility perspectives. *Scandinavian Journal of Management*, *40*(4), 101369. https://doi.org/10.1016/j.scaman.2024.101369

Ma, R., Cherian, J., Tsai, W.-H., Sial, M. S., Hou, L., & Álvarez-Otero, S. (2021). The Relationship of Corporate Social Responsibility on Digital Platforms, Electronic Word-of-Mouth, and Consumer-Company Identification: An Application of Social Identity Theory. *Sustainability*, *13*(9), 4700. https://doi.org/10.3390/su13094700

Ma, Y., Guo, X., Su, W., & Fu, G. (2024). The Evolution of Price Discrimination in E-Commerce Platform Trading: A Perspective of Platform Corporate Social Responsibility. *Journal of Theoretical and Applied Electronic Commerce Research*, *19*(3), 1907–1921. https://doi.org/10.3390/jtaer19030094

Mayer, D. M., Nurmohamed, S., Treviño, L. K., Shapiro, D. L., & Schminke, M. (2013a). Encouraging employees to report unethical conduct internally: It takes a village. *Organizational Behavior and Human Decision Processes*, *121*(1), 89–103. https://doi.org/10.1016/j.obhdp.2013.01.002

Mayer, D. M., Nurmohamed, S., Treviño, L. K., Shapiro, D. L., & Schminke, M. (2013b). Encouraging employees to report unethical conduct internally: It takes a village. *Organizational Behavior and Human Decision Processes*, *121*(1), 89–103. https://doi.org/10.1016/j.obhdp.2013.01.002

McBride, C., Costello, N., Ambwani, S., Wilhite, B., & Austin, S. B. (2019). Digital Manipulation of Images of Models' Appearance in Advertising: Strategies for Action Through Law and Corporate Social Responsibility Incentives to Protect Public Health. *American Journal of Law & Medicine*, *45*(1), 7–31. https://doi.org/10.1177/0098858819849990

McGlynn, J., & Richardson, B. K. (2014). Private Support, Public Alienation: Whistle-Blowers and the Paradox of Social Support. *Western Journal of Communication*, *78*(2), 213–237. https://doi.org/10.1080/10570314.2013.807436

McIntosh, T., Higgs, C., Turner, M., Partlow, P., Steele, L., MacDougall, A. E., Connelly, S., & Mumford, M. D. (2019). To Whistleblow or Not to Whistleblow: Affective and Cognitive Differences in Reporting Peers and Advisors. *Science and Engineering Ethics*, *25*(1), 171–210. https://doi.org/10.1007/s11948-017-9974-3

McLain, D. L., & Keenan, J. P. (1999). Risk, information, and the decision about response to wrongdoing in an organization. *Journal of Business Ethics*, *19*(3), 255–271. https://doi.org/10.1023/A:1006168301995

McVea, J. F., & Freeman, R. E. (2005). A Names-and-Faces Approach to Stakeholder Management: How Focusing on Stakeholders as Individuals Can Bring Ethics and

Entrepreneurial Strategy Together. *Journal of Management Inquiry*, *14*(1), 57–69. https://doi.org/10.1177/1056492604270799

McWilliams, A., & Siegel, D. (2001). Corporate Social Responsibility: A Theory of the Firm Perspective. *Academy of Management Review*, *26*(1), 117–127.

Mechtenberg, L., Muehlheusser, G., & Roider, A. (2020). Whistleblower protection: Theory and experimental evidence. *European Economic Review*, *126*, 103447. https://doi.org/10.1016/j.euroecorev.2020.103447

Merbecks, U. (2024). Corporate digital responsibility (CDR) in Germany: Background and first empirical evidence from DAX 30 companies in 2020. *Journal of Business Economics*, *94*(7–8), 1025–1049. https://doi.org/10.1007/s11573-023-01148-6

Merguei, N., Strobel, M., & Vostroknutov, A. (2022). Moral opportunism as a consequence of decision making under uncertainty. *Journal of Economic Behavior & Organization*, *197*, 624–642. https://doi.org/10.1016/j.jebo.2022.03.020

Mesa-Vázquez, E., Rodriguez-Lara, I., & Urbano, A. (2021). Standard vs random dictator games: On the effects of role uncertainty and framing on generosity. *Economics Letters*, *206*, 109981. https://doi.org/10.1016/j.econlet.2021.109981

Mesmer-Magnus, J. R., & Viswesvaran, C. (2005). Whistleblowing in organizations: An examination of correlates of whistleblowing intentions, actions, and retaliation. *Journal of Business Ethics*, *62*(3), 277–297. https://doi.org/10.1007/s10551-005-0849-1

Miceli, M. P., Dozier, J. B., & Near, J. P. (1991). Blowing the whistle on data fudging: A controlled field experiment. *Journal of Applied Social Psychology*, *21*(4), 271–295.

Miceli, M. P., & Near, J. P. (1984). The Relationships among Beliefs, Organizational Position, and Whistle-Blowing Status: A Discriminant Analysis. *The Academy of Management Journal*, *27*(4), 687–705.

Miceli, M. P., & Near, J. P. (1992). *Blowing the whistle: The organizational and legal implications for companies and employees*. Lexington Books.

Miceli, M. P., Near, J. P., Rehg, M. T., & van Scotter, J. R. (2012). Predicting employee reactions to perceived organizational wrongdoing: Demoralization, justice, proactive personality, and whistle-blowing. In *Human Relations* (Bd. 65, Nummer 8). https://doi.org/10.1177/0018726712447004

Miethe, T. D., & Rothschild, J. (1994). Whistleblowing and the control of organizational misconduct. *Sociological Inquiry*, *64*(3), 322–347.

Mihale-Wilson, C., Hinz, O., Van Der Aalst, W., & Weinhardt, C. (2022). Corporate Digital Responsibility: Relevance and Opportunities for Business and Information Systems Engineering. *Business & Information Systems Engineering*, *64*(2), 127–132. https://doi.org/10.1007/s12599-022-00746-y

Miles, S. (2017). Stakeholder Theory Classification: A Theoretical and Empirical Evaluation of Definitions. *Journal of Business Ethics*, *142*(3), 437–459. https://doi.org/10.1007/s10551-015-2741-y

Miller, D. T., & Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, *67*, 339–361.

Mir Djawadi, B., Plaß, S., & Schäfers, S. (2023). Peers: Powerful or Negligible? A Systematic Review on Peer Factors and Internal Whistleblowing. In C. Gabbioneta, M. Clemente, & R. Greenwood (Hrsg.), *Organizational Wrongdoing as the "Foundational" Grand Challenge: Consequences and Impact* (S. 73–100). Emerald Publishing Limited.

Mir Djawadi, B., Plaß, S., & Schäfers, S. (2025). *"I don't believe that you believe what I believe": An experiment on misperceptions of social norms and whistleblowing*. https://dx.doi.org/10.2139/ssrn.4868252

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 2053951716679679. https://doi.org/10.1177/2053951716679679

Mogaji, E., Wirtz, J., Belk, R. W., & Dwivedi, Y. K. (2023). Immersive time (ImT): Conceptualizing time spent in the metaverse. *International Journal of Information Management*, *72*, 102659. https://doi.org/10.1016/j.ijinfomgt.2023.102659

Montiel, I. (2008). Corporate Social Responsibility and Corporate Sustainability: Separate Pasts, Common Futures. *Organization & Environment*, *21*(3), 245–269. https://doi.org/10.1177/1086026608321329

Montiel, I., Gallo, P. J., & Antolin-Lopez, R. (2020). What on Earth Should Managers Learn About Corporate Sustainability? A Threshold Concept Approach. *Journal of Business Ethics*, *162*(4), 857–880. https://doi.org/10.1007/s10551-019-04361-y

Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true North, and how we can fix it. *Research in Organizational Behavior*, *33*, 53–77. https://doi.org/10.1016/j.riob.2013.08.001

Muehlheusser, G., & Roider, A. (2008). Black sheep and walls of silence. *Journal of Economic Behavior and Organization*, *65*(3–4), 387–408.

Mueller, B. (2022). Corporate Digital Responsibility. *Business & Information Systems Engineering*, *64*(5), 689–700. https://doi.org/10.1007/s12599-022-00760-0

Mukhuty, S., Upadhyay, A., & Rothwell, H. (2022). Strategic sustainable development of Industry 4.0 through the lens of social responsibility: The role of human resource practices. *Business Strategy and the Environment*, *31*(5), 2068–2081. https://doi.org/10.1002/bse.3008

Munsch, C. L., Ridgeway, C. L., & Williams, J. C. (2014). Pluralistic Ignorance and the Flexibility Bias: Understanding and Mitigating Flextime and Flexplace Bias at Work. *Work and Occupations*, *41*(1), 40–62. https://doi.org/10.1177/0730888413515894

Murphy, A. R. (2021). *Conscience and community*. Penn State University Press.

Nagano, A. (2023). Institutional Values and CDR for Green ITC: Transforming E-waste to Ethical Gems. *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 1–6. https://doi.org/10.1109/WINCOM59760.2023.10322959

Napoli, F. (2023). Corporate Digital Responsibility: A Board of Directors May Encourage the Environmentally Responsible Use of Digital Technology and Data: Empirical Evidence from Italian Publicly Listed Companies. *Sustainability*, *15*(3), 2539. https://doi.org/10.3390/su15032539

Nayal, K., Kumar, S., Raut, R. D., Queiroz, M. M., Priyadarshinee, P., & Narkhede, B. E. (2021). Supply chain firm performance in circular economy and digital era to achieve sustainable development goals. *Business Strategy and the Environment*, *31*(3), 1058–1073. https://doi.org/10.1002/bse.2935

Near, J. P., & Miceli, M. P. (1985). Organizational Dissidence: The Case of Whistle-Blowing.

Near, J. P., & Miceli, M. P. (1995). Effective Whistle-Blowing. *The Academy of Management Review*, *20*(3), 679–708.

Near, J. P., Rehg, M. T., Van Scotter, J. R., & Miceli, M. P. (2004). Does Type of Wrongdoing Affect the Whistle-Blowing Process? *Business Ethics Quarterly*, *14*(2), 219–242. https://doi.org/10.5840/beq200414210

Neligan, A., Baumgartner, R. J., Geissdoerfer, M., & Schöggl, J. (2023). Circular disruption: Digitalisation as a driver of circular economy business models. *Business Strategy and the Environment*, *32*(3), 1175–1188. https://doi.org/10.1002/bse.3100

Neri, A., Negri, M., Cagno, E., Kumar, V., & Garza-Reyes, J. A. (2023). What digital-enabled dynamic capabilities support the circular economy? A multiple case study approach. *Business Strategy and the Environment*, *32*(7), 5083–5101. https://doi.org/10.1002/bse.3409

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems*, *24*(1), 3–14. https://doi.org/10.1016/j.jsis.2015.02.001

Nicholls, A. R., Fairs, L. R. W., Toner, J., Jones, L., Mantis, C., Barkoukis, V., Perry, J. L., Micle, A. V., Theodorou, N. C., Shakhverdieva, S., Stoicescu, M., Vesic, M. V., Dikic, N., Andjelkovic, M., Grimau, E. G., Amigo, J. A., & Schomöller, A. (2021). Snitches Get Stitches and End Up in Ditches: A Systematic Review of the Factors Associated With Whistleblowing Intentions. *Frontiers in Psychology*, *12*(October). https://doi.org/10.3389/fpsyg.2021.631538

Nie, J., Jian, X., Xu, J., Xu, N., Jiang, T., & Yu, Y. (2024). The effect of corporate social responsibility practices on digital transformation in China: A resource-based view. *Economic Analysis and Policy*, *82*, 1–15. https://doi.org/10.1016/j.eap.2024.02.027

Niehoff, S. (2022). Aligning digitalisation and sustainable development? Evidence from the analysis of worldviews in sustainability reports. *Business Strategy and the Environment*, *31*(5), 2546–2567. https://doi.org/10.1002/bse.3043

O'Fallon, M. J., & Butterfield, K. D. (2005). A Review of The Empirical Ethical Decision-Making Literature: 1996–2003. *Journal of Business Ethics*, *59*(4), 375–413. https://doi.org/10.1007/s10551-005-2929-7

Oh, H., Chung, M. H. O., & Labianca, G. (2004). Group social capital and group effectiveness: The role of informal socializing ties. *Academy of Management Journal*, *47*(6), 860–875. https://doi.org/10.5465/20159627

Okazaki, S., Plangger, K., West, D., & Menéndez, H. D. (2020). Exploring digital corporate social responsibility communications on Twitter. *Journal of Business Research*, *117*, 675–682. https://doi.org/10.1016/j.jbusres.2019.09.006

Okorie, O., Russell, J., Cherrington, R., Fisher, O., & Charnley, F. (2023). Digital transformation and the circular economy: Creating a competitive advantage from the transition towards Net Zero Manufacturing. *Resources, Conservation and Recycling*, *189*, 106756. https://doi.org/10.1016/j.resconrec.2022.106756

O'Leary, C., & Pangemanan, G. (2007). The effect of groupwork on ethical decision-making of accountancy students. *Journal of Business Ethics*, *75*(3), 215–228. https://doi.org/10.1007/s10551-006-9248-5

Olesen, T. (2019). The Politics of Whistleblowing in Digitalized Societies. *Politics & Society*, *47*(2), 277–297.

Orbik, Z., & Zozuľaková, V. (2019). Corporate Social and Digital Responsibility. *Management Systems in Production Engineering*, *27*(2), 79–83. https://doi.org/10.1515/mspe-2019-0013

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington Books.

Özturan, P., & Grinstein, A. (2022). Impact of Global Brand Chief Marketing Officers' Corporate Social Responsibility and Sociopolitical Activism Communication on Twitter. *Journal of International Marketing*, *30*(3), 72–82. https://doi.org/10.1177/1069031X221104077

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Paluch, S., Pitardi, V., & Kunz, W. H. (2024). The ethical edge: Understanding corporate digital responsibility in organizations. *Organizational Dynamics*, *53*(2), 101058. https://doi.org/10.1016/j.orgdyn.2024.101058

Pan, S. L., & Nishant, R. (2023). Artificial intelligence for digital sustainability: An insight into domain-specific research and future directions. *International Journal of Information Management*, *72*, 102668. https://doi.org/10.1016/j.ijinfomgt.2023.102668

Pan, X., Oh, K.-S., & Wang, M. (2021). Strategic Orientation, Digital Capabilities, and New Product Development in Emerging Market Firms: The Moderating Role of Corporate Social Responsibility. *Sustainability*, *13*(22), 12703. https://doi.org/10.3390/su132212703

Panizza, F., Dimant, E., Kimbrough, E. O., & Vostroknutov, A. (2023). Measuring Norm Pluralism and Tolerance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4649792

Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., & Krogstie, J. (2023). Responsible Digital Transformation for a Sustainable Society. *Information Systems Frontiers*, *25*(3), 945–953. https://doi.org/10.1007/s10796-023-10406-5

Park, H., Kim, Y., & Popelish, B. (2021). Corporate Social Responsibility and Stakeholder Engagement: A Content Analysis of PRSA Silver Anvil Award-Winning CSR Campaigns. *Journal of Sustainability Research*, *3*(3). https://doi.org/10.20900/jsr20210015

Park, S. H., Gonzalez-Perez, M. A., & Floriani, D. E. (Hrsg.). (2021). *The Palgrave Handbook of Corporate Sustainability in the Digital Era*. Springer International Publishing. https://doi.org/10.1007/978-3-030-42412-1

Parmar, B. L., Freeman, R. E., & Harrison, J. S. (2010). Stakeholder Theory: The State of the Art. *Academy of Management Annals*, *4*(1), 403–445.

Parmentola, A., Petrillo, A., Tutore, I., & De Felice, F. (2022). Is blockchain able to enhance environmental sustainability? A systematic review and research agenda from the perspective of Sustainable Development Goals (SDGs). *Business Strategy and the Environment*, *31*(1), 194–217. https://doi.org/10.1002/bse.2882

Patil, A., Dwivedi, A., Abdul Moktadir, Md., & Lakshay. (2023). Big data-Industry 4.0 readiness factors for sustainable supply chain management: Towards circularity. *Computers & Industrial Engineering*, *178*, 109109. https://doi.org/10.1016/j.cie.2023.109109

Paul, J., Ueno, A., Dennis, C., Alamanos, E., Curtis, L., Foroudi, P., Kacprzak, A., Kunz, W. H., Liu, J., Marvi, R., Nair, S. L. S., Ozdemir, O., Pantano, E., Papadopoulos, T., Petit, O., Tyagi, S., & Wirtz, J. (2024). Digital transformation: A multidisciplinary perspective and future research agenda. *International Journal of Consumer Studies*, *48*(2), e13015. https://doi.org/10.1111/ijcs.13015

Pauliuk, S., Koslowski, M., Madhu, K., Schulte, S., & Kilchert, S. (2022). Co-design of digital transformation and sustainable development strategies—What socio-metabolic and industrial ecology research can contribute. *Journal of Cleaner Production*, *343*, 1–12. https://doi.org/10.1016/j.jclepro.2022.130997

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Pelters, E. (2021). Corporate Digital Responsibility—Understanding and Applying. In T. A. Herberger & J. J. Dötsch (Hrsg.), *Digitalization, Digital Transformation and Sustainability in the Global Economy* (S. 71–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-77340-3_6

Peng, H., Zhang, Q., & Zhang, Z. (2024). Does Information Source Matter? Corporate Reputation Management during Negative Social Responsibility Events. *Journal of Theoretical and Applied Electronic Commerce Research*, *19*(4), 2747–2764. https://doi.org/10.3390/jtaer19040132

Pershing, J. L. (2002). Whom to betray? Self-regulation of occupational misconduct at the United States Naval Academy. *Deviant Behavior*, *23*(2), 149–175. https://doi.org/10.1080/016396202753424538

Peshkova, O. A. (2022). Digital Footprint Analysis Technology: Some Aspects of Its Application in Recruitment. In S. I. Ashmarina, V. V. Mantulenko, & M. Vochozka (Hrsg.), *Proceedings of the International Scientific Conference "Smart Nations: Global Trends In The Digital Economy"* (Bd. 397, S. 368–375). Springer International Publishing. https://doi.org/10.1007/978-3-030-94873-3_46

Pînzaru, F., Dima, A. M., Zbuchea, A., & Veres, Z. (2022). Adopting Sustainability and Digital Transformation in Business in Romania: A Multifaceted Approach in the Context of the Just Transition. *Www.Amfiteatrueconomic.Ro*, *24*(59), 28. https://doi.org/10.24818/EA/2022/59/28

Pranugrahaning, A., Donovan, J. D., Topple, C., & Masli, E. K. (2021). Corporate sustainability assessments: A systematic literature review and conceptual framework. *Journal of Cleaner Production*, *295*, 126385. https://doi.org/10.1016/j.jclepro.2021.126385

Prihadyanti, D., & Aziz, S. A. (2023). Indonesia toward sustainable agriculture–Do technology-based start-ups play a crucial role? *Business Strategy & Development*, *6*(2), 140–157.

Prisco, A., Sepe, F., Nanu, L., & Tani, M. (2024). Exploring food delivery app adoption: Corporate social responsibility and perceived product risk's influence. *Corporate Social Responsibility and Environmental Management*, csr.3041. https://doi.org/10.1002/csr.3041

Pristl, A., Kilian, S., & Mann, A. (2021). When does a social norm catch the worm? Disentangling social normative influences on sustainable consumption behaviour. *Journal of Consumer Behaviour*, *20*(3), 635–654. https://doi.org/10.1002/cb.1890

Rangel-Pérez, C., Fernández, M., & López, B. (2023). Study on the strategic influence of corporate social responsibility in the world's most digitised banks. *Journal of Open Innovation: Technology, Market, and Complexity*, *9*(1), 100029. https://doi.org/10.1016/j.joitmc.2023.100029

Raworth, K. (2017). *Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist*. Chelsea Green Publishing. https://books.google.de/books?id=7A4lDgAAQBAJ

Raworth, K. (2018). *Doughnut economics: Seven ways to think like a 21st century economist.* Chelsea Green Publishing.

Raymond, S., Beddoe, L., & Staniforth, B. (2017). Social workers' experiences with whistleblowing: To speak or not to speak? *Aotearoa New Zealand Social Work*, *29*(3), 17–29. https://doi.org/10.11157/anzswj-vol29iss3id305

Rejeb, A., & Rejeb, K. (2020). Blockchain and supply chain sustainability. *Logforum*, *16*(3), 363–372. https://doi.org/10.17270/J.LOG.2020.467

Rennie, S. C., & Crosby, J. R. (2002). Students' perceptions of whistle blowing: Implications for self-regulation. A questionnaire and focus group survey. *Medical Education*, *36*(2), 173–179. https://doi.org/10.1046/j.1365-2923.2002.01137.x

Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). *Postconventional moral thinking: A Neo-Kohlbergian approach.* Psychology Press.

Reuben, E., & Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, *77*(1), 122–137. https://doi.org/10.1016/j.geb.2012.10.001

Reuben, E., & Stephenson, M. (2013). Nobody likes a rat: On the willingness to report lies and the consequences thereof. *Journal of Economic Behavior and Organization*, *93*, 384–391. https://doi.org/10.1016/j.jebo.2013.03.028

Ribeiro, G. C., Oliveira, K. K. S., & Souza, R. A. C. (2024). DSI Strategy Canvas: Modelling the Digital Social Innovation Strategy. *Journal of Social Entrepreneurship*, *15*(2), 630–658. https://doi.org/10.1080/19420676.2021.1987971

Rimal, R. N., & Real, K. (2003). Understanding the Influence of Perceived Norms on Behaviors. *Communication Theory*, *13*(2), 184–203.

Rodriguez-Gomez, S., Arco-Castro, M. L., Lopez-Perez, M. V., & Rodríguez-Ariza, L. (2020). Where does csr come from and where does it go? A review of the state of the art. *Administrative Sciences*, *10*(3). https://doi.org/10.3390/admsci10030060

Rogers, S. (2012). Britain's top 1,000 charities ranked by donations. Who raises the most money? *The Guardian*. https://www.theguardian.com/news/datablog/2012/apr/24/top-1000-charities-donations-britain

Rothschild, J. (2008). Freedom of Speech Denied, Dignity Assaulted: What the Whistleblowers Experience in the US. *Current Sociology*, *56*(6), 884–903. https://doi.org/10.1177/0011392108095344

Rothschild, J., & Miethe, T. D. (1999). Whistle-Blower Disclosures and Management Retaliation: The Battle to Control Information about Organization Corruption. *Work and Occupations*, *26*(1), 107–128. https://doi.org/10.1177/0730888499026001006

Rothwell, G. R., & Baldwin, J. N. (2007). Ethical climate theory, whistle-blowing, and the code of silence in police agencies in the State of Georgia. *Journal of Business Ethics*, *70*(4), 341–361. https://doi.org/10.1007/s10551-006-9114-5

Rousseau, D., Manning, J., & Denyer, D. (2008). Science: Assembling the Field ' s Full Weight of Scientific Knowledge Through Syntheses. *AIM Research Working Paper Series*, *67*(8), 1–78.

Rugeviciute, A. (2023). Analysis and modelling of socio-environmental impacts of Corporate Digital Responsibility. In *CEUR Workshop Proceedings* (Bd. 3562, S. 18–27). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85180011985&partnerID=40&md5=e71c4362c88e78b814aeef661c00bc37

Rusch, M., Schöggl, J., & Baumgartner, R. J. (2022). Application of digital technologies for sustainable product management in a circular economy: A review. *Business Strategy and the Environment*, *32*(3), 1159–1174. https://doi.org/10.1002/bse.3099

Sætra, H. S. (2023). The AI ESG protocol: Evaluating and disclosing the environment, social, and governance implications of artificial intelligence capabilities, assets, and activities. *Sustainable Development*, *31*(2), 1027–1037. https://doi.org/10.1002/sd.2438

Sahu, A., Agrawal, S., & Kumar, G. (2022). Integrating Industry 4.0 and circular economy: A review. *Journal of Enterprise Information Management*, *35*(3), 885–917. https://doi.org/10.1108/JEIM-11-2020-0465

Sahu, A. K., Sahu, N. K., & Sahu, A. K. (2023). Laminating STRATH block chain technology-SWOT architectures to endure business strategy between digital transformation, firms and supply chains capabilities for sustainability. *Journal of Cleaner Production*, *383*, 135531. https://doi.org/10.1016/j.jclepro.2022.135531

Salancik, G. R., & Pfeffer, J. (1978). A Social Information Processing Approach to Job Attitudes and Task Design. *Administrative Science Quarterly*, *23*(2), 224. https://doi.org/10.2307/2392563

Santarius, T. (2016). Investigating meso-economic rebound effects: Production-side effects and feedback loops between the micro and macro level. *Journal of Cleaner Production*, *134*, 406–413. https://doi.org/10.1016/j.jclepro.2015.09.055

Santarius, T., & Wagner, J. (2023). Digitalization and sustainability: A systematic literature analysis of *ICT for Sustainability* research. *GAIA - Ecological Perspectives for Science and Society*, *32*(1), 21–32. https://doi.org/10.14512/gaia.32.S1.5

Sargent, R. H., & Newman, L. S. (2021). Pluralistic Ignorance Research in Psychology: A Scoping Review of Topic and Method Variation and Directions for Future Research. *Review of General Psychology*, *25*(2), 163–184.

Scarpi, D., & Pantano, E. (2024). "With great power comes great responsibility": Exploring the role of Corporate Digital Responsibility (CDR) for Artificial Intelligence Responsibility in Retail Service Automation (AIRRSA). *Organizational Dynamics*, *53*(2), 101030. https://doi.org/10.1016/j.orgdyn.2024.101030

Schlütter, D., Schätzlein, L., Hahn, R., & Waldner, C. (2024). Missing the Impact in Impact Investing Research – A Systematic Review and Critical Reflection of the Literature. *Journal of Management Studies*, *61*(6), 2694–2718. https://doi.org/10.1111/joms.12978

Schneider, G. (2022). Framing Accountability in Business-to-Government Data Sharing: The Gap Filling Role of Businesses' Corporate Digital Responsibility. *European Business Law Review*, *33*(6).

Schons, L., & Steinmeier, M. (2016). Walk the Talk? How Symbolic and Substantive CSR Actions Affect Firm Performance Depending on Stakeholder Proximity. *Corporate Social Responsibility and Environmental Management*, *23*(6), 358–372. https://doi.org/10.1002/csr.1381

Schrödter, A., & Weißenberger, B. E. (2024). The institutionalization of digital compliance. *Management Decision*. https://doi.org/10.1108/MD-03-2024-0498

Schwepker, C. H. (1999). Research Note: The Relationship between Ethical Conflict, Organizational Commitment and Turnover Intentions in the Salesforce. *Sales Management*.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationale Verhaltens im Rahmen eines Oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung*.

Shen, J., & Benson, J. (2016). When CSR Is a Social Norm: How Socially Responsible Human Resource Management Affects Employee Work Behavior. *Journal of Management*, *42*(6), 1723–1746. https://doi.org/10.1177/0149206314522300

Shen, Q., Du, Y., & Huang, J. (2024). Consumer Citizenship Behavior in Online/Offline Shopping Contexts: Differential Impact of Consumer Perceived Value and Perceived Corporate Social Responsibility. *Sustainability*, *16*(7), 2968. https://doi.org/10.3390/su16072968

Shestakova, I. (2024). The Era of Digital Transition in the Prism of the Existential Threat of Job Loss: Corporate Social Responsibility. *Sustainability*, *16*(18), 8019. https://doi.org/10.3390/su16188019

Shkalenko, A. V., & Nazarenko, A. V. (2024). Integration of AI and IoT into Corporate Social Responsibility Strategies for Financial Risk Management and Sustainable Development. *Risks*, *12*(6), 87. https://doi.org/10.3390/risks12060087

Sidaoui, K., Mahr, D., & Odekerken-Schröder, G. (2024). Generative AI in Responsible Conversational Agent Integration: Guidelines for Service Managers. *Organizational Dynamics*, *53*(2), 101045. https://doi.org/10.1016/j.orgdyn.2024.101045

Skivenes, M., & Trygstad, S. C. (2016). Whistleblowing in Local Government: An Empirical Study of Contact Patterns and Whistleblowing in 20 Norwegian Municipalities. *Scandinavian Political Studies*, *39*(3), 264–289. https://doi.org/10.1111/1467-9477.12066

Spoelma, T. M., Chawla, N., & Ellis, A. P. J. (2021). If You Can't Join 'Em, Report 'Em: A Model of Ostracism and Whistleblowing in Teams. *Journal of Business Ethics*, *173*(2), 345–363. https://doi.org/10.1007/s10551-020-04563-9

Srivetbodee, S., & Igel, B. (2021). Digital Technology Adoption in Agriculture: Success Factors, Obstacles and Impact on Corporate Social Responsibility Performance in Thailand's Smart Farming Projects. *Thammasat Review*, *24*, 149170. https://doi.org/10.14456/TUREVIEW.2021.22

Stahl, B. C. (2024). From Corporate Digital Responsibility to Responsible Digital Ecosystems. *Sustainability*, *16*(12), 4972. https://doi.org/10.3390/su16124972

Stock, C., Hossinger, S., Werner, A., Schell, S., & Soluk, J. (2022). Corporate social responsibility as a driver of digital innovation in SMEs: The mediation effect of absorptive capacity. *International Journal of Entrepreneurial Venturing*, *14*(4–5), 571–601.

Strand, R. (2023). Strategic Leadership of Corporate Sustainability. *Journal of Business Ethics*, *123*(4), 687–706. https://doi.org/10.1007/s10551-013-2017-3

Suchacka, M. (2019). Corporate Digital Responsibility New Challenges to the Social Sciences. *International Journal of Research in E-Learning*, *5*(1), 5–20. https://doi.org/10.31261/IJREL.2019.5.1.01

Suchacka, M. (2020). Corporate Digital Responsibility—A New Dimension of the Human—Technology Relations. *System Safety: Human - Technical Facility - Environment*, *2*(1), 1–8. https://doi.org/10.2478/czoto-2020-0001

Sun, Z., Wang, W., Wang, W., & Sun, X. (2024). How does digital transformation affect corporate social responsibility performance? From the dual perspective of internal drive and external governance. *Corporate Social Responsibility and Environmental Management*, *31*(2), 1156–1176. https://doi.org/10.1002/csr.2615

Tajfel, H., & Turner, J. C. (2004). The Social Identity Theory of Intergroup Behavior. In *Political psychology: Key readings* (S. 276–293). J. T. Jost & J. Sidanius.

Tarafdar, M., Tu, Q., Ragu-Nathan, B. S., & Ragu-Nathan, T. (2007). The impact of technostress on role stress and productivity. *Journal of Management Information Systems*, *24*(1), 301–328.

Tasleem, M., Khan, N., & Nisar, A. (2019). Impact of technology management on corporate sustainability performance: The mediating role of TQM. *International Journal of Quality & Reliability Management*, *36*(9), 1574–1599. https://doi.org/10.1108/IJQRM-01-2018-0017

Taylor, E. Z., & Curtis, M. B. (2010). An Examination of the Layers of Workplace Influences in Ethical Judgments: Whistleblowing Likelihood and Perseverance in Public Accounting. *Journal of Business Ethics*, *93*(1), 21–37. https://doi.org/10.1007/s10551-009-0179-9

Taylor, E. Z., & Curtis, M. B. (2013). Whistleblowing in audit firms: Organizational response and power distance. *Behavioral Research in Accounting*, *25*(2), 21–43. https://doi.org/10.2308/bria-50415

Taylor, J. (2018). Internal Whistle-Blowing in the Public Service: A Matter of Trust. *Public Administration Review*, *78*(5), 717–726. https://doi.org/10.1111/puar.12946

Tenbrunsel, A. E., Smith-Crowe, K., & Umphress, E. E. (2003). Building houses on rocks: The role of the ethical infrastructure in organizations. *Social Justice Research*, *16*(3), 285–307. https://doi.org/10.1023/A:1025992813613

Teo, H., & Caspersz, D. (2011). Dissenting discourse: Exploring alternatives to the whistleblowing/silence dichotomy. *Journal of Business Ethics*, *104*, 237–249.

Tettamanzi, P., Venturini, G., & Murgolo, M. (2022). Sustainability and Financial Accounting: A Critical Review on the ESG Dynamics. *Environmental Science and Pollution Research*, *29*(11), 16758–16761. https://doi.org/10.1007/s11356-022-18596-2

Thorpe, R., Holt, R., Macpherson, A., & Pittaway, L. (2005a). Using knowledge within small and medium-sized firms: A systematic review of the evidence. *International Journal of Management Reviews*, *7*, 257–281.

Thorpe, R., Holt, R., Macpherson, A., & Pittaway, L. (2005b). Using knowledge within small and medium-sized firms: A systematic review of the evidence. *International Journal of Management Reviews*, *7*(4), 257–281. https://doi.org/10.1111/j.1468-2370.2005.00116.x

Thuong, D. T. M. (2024). The Joint Effect of Corporate Social Responsibility and Digital Transformation on Bank Performance: Evidence from Vietnam'S Banking Sector. *Journal of Logistics, Informatics and Service Science*, *11*(12). https://doi.org/10.33168/JLISS.2024.1216

Torrent-Sellens, J., Ficapal-Cusí, P., & Enache-Zegheru, M. (2023). Boosting environmental management: The mediating role of Industry 4.0 between environmental assets and economic and social firm performance. *Business Strategy and the Environment*, *32*(1), 753–768. https://doi.org/10.1002/bse.3173

Tóth, Z., & Blut, M. (2024). Ethical compass: The need for Corporate Digital Responsibility in the use of Artificial Intelligence in financial services. *Organizational Dynamics*, *53*(2), 101041. https://doi.org/10.1016/j.orgdyn.2024.101041

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Treviño, L. K. (1986). Ethical Decision Making in Organizations: A Person-Situation Interactionist Model. *Academy of Management Review*, *11*(3), 601–617.

Treviño, L. K., Butterfield, K. D., & McCabe, D. L. (1998). The Ethical Context in Organizations: Influences on Employee Attitudes and Behaviors. *Business Ethics Quarterly*, *8*(3), 447–476. https://doi.org/10.2307/3857431

Treviño, L. K., & Victor, B. (1992). Peer Reporting of Unethical Behavior: A Social Context Perspective. *The Academy of Management Journal*, *35*(1), 38–64.

Trier, M., Kundisch, D., Beverungen, D., Müller, O., Schryen, G., Mirbabaie, M., & Trang, S. (2023). Digital Responsibility: A Multilevel Framework for Responsible Digitalization. *Business & Information Systems Engineering*, *65*(4), 463–474. https://doi.org/10.1007/s12599-023-00822-x

Trittin-Ulbrich, H., & Böckel, A. (2022). Institutional entrepreneurship for responsible digital innovation: The case of corporate digital responsibility. *Creativity and Innovation Management*, *31*(3), 447–459. https://doi.org/10.1111/caim.12513

Trongmateerut, P., & Sweeney, J. T. (2013). The influence of subjective norms on whistle-blowing: A cross-cultural investigation. *Journal of Business Ethics*, *112*, 437–451.

Tsahuridu, E. E., & Vandekerckhove, W. (2008). Organisational whistleblowing policies: Making employees responsible or liable? *Journal of Business Ethics*, *82*(1), 107–118. https://doi.org/10.1007/s10551-007-9565-3

Tuyen, B. Q., Phuong Anh, D. V., Mai, N. P., & Long, T. Q. (2023). Does corporate engagement in social responsibility affect firm innovation? The mediating role of digital transformation. *International Review of Economics & Finance*, *84*, 292–303. https://doi.org/10.1016/j.iref.2022.11.005

Ukko, J., Nasiri, M., Saunila, M., & Rantala, T. (2019). Sustainability strategy as a moderator in the relationship between digital business strategy and financial performance. *Journal of Cleaner Production*, *236*, 117626. https://doi.org/10.1016/j.jclepro.2019.117626

Vadera, A. K., Aguilera, R. V., & Caza, B. B. (2009). Making Sense of Whistle-Blowing's Antecedents: Learning from Research on Identity and Ethics Programs. *Business Ethics Quarterly*, *19*(4), 553–586. https://doi.org/10.5840/beq200919432

Van Der Linden, S. (2015). The social-psychological determinants of climate change risk perceptions: Towards a comprehensive model. *Journal of Environmental Psychology*, *41*, 112–124. https://doi.org/10.1016/j.jenvp.2014.11.012

Van Der Merwe, J., & Al Achkar, Z. (2022). Data responsibility, corporate social responsibility, and corporate digital responsibility. *Data & Policy*, *4*(2), 12. https://doi.org/10.1017/dap.2022.2

Van Marrewijk, M. (2003). Concepts and Definitions of CSR and Corporate Sustainability: Between Agency and Communion. In A. C. Michalos & D. C. Poff (Hrsg.), *Citation Classics from the Journal of Business Ethics* (S. 641–655). Springer Netherlands. https://doi.org/10.1007/978-94-007-4126-3_32

Van Portfliet, M. (2020). Resistance Will Be Futile? The Stigmatization (or Not) of Whistleblowers. *Journal of Business Ethics*, *175*(3), 451–464. https://doi.org/10.1007/s10551-020-04673-4

Van Zanten, J. A., & Van Tulder, R. (2021). Improving companies' impacts on sustainable development: A nexus approach to the SDGS. *Business Strategy and the Environment*, *30*(8), 3703–3720. https://doi.org/10.1002/bse.2835

Vandekerckhove, W., Brown, A. J., & Tsahurid, E. (2014). Managerial responsiveness to whistleblowing: Expanding the research horizon. In *International handbook on whistleblowing research* (S. 298–328). Edward Elgar Publishing.

Veselý, Š. (2015). Elicitation of normative and fairness judgments: Do incentives matter? *Judgment and Decision Making*, *10*(2), 191–197. https://doi.org/10.1017/S1930297500003958

Vítová, B. (2022). Unfair Commercial Practices and Corporate Social Responsibility (CSR) In the Digital World. *European Studies*, *9*(2), 197–212. https://doi.org/10.2478/eustu-2022-0020

Vo Thai, H., Hue, T. H. H., Chen, P., & Tran, M. (2024). Unraveling the influence of human capital and stakeholder engagement on corporate digital responsibility: Implications for firm performance in Southeast Asia enterprises. *Corporate Social Responsibility and Environmental Management*, *31*(3), 1934–1958. https://doi.org/10.1002/csr.2662

Volchek, R., Moskaliuk, H., Halan, L., & Dancheva, O. (2024). Implementation of corporate social responsibility in the context of integration with the enterprise management information system. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, *1*, 154–161. https://doi.org/10.33271/nvngu/2024-1/154

Volkov, V. R., & Sidorenko, E. L. (2022). Digital Platforms and Issues of Corporate Criminal Responsibility, Self-regulation. In S. I. Ashmarina, V. V. Mantulenko, & M. Vochozka (Hrsg.), *Proceedings of the International Scientific Conference "Smart Nations: Global Trends In The Digital Economy"* (Bd. 397, S. 3–9). Springer International Publishing. https://doi.org/10.1007/978-3-030-94873-3_1

Waddock, S. A., & Graves, S. B. (1997). The Corporate Social Performance-Financial Performance Link. *Strategic Management Journal*, *18*(4), 303–319.

Wagner, T., Lutz, R. J., & Weitz, B. A. (2009). Corporate Hypocrisy: Overcoming the Threat of Inconsistent Corporate Social Responsibility Perceptions. *Journal of Marketing*, *73*(6), 77–91.

Walkowitz, G. (2021). Dictator game variants with probabilistic (and cost-saving) payoffs: A systematic test. *Journal of Economic Psychology*, *85*, 102387. https://doi.org/10.1016/j.joep.2021.102387

Wang, H. C., He, J., & Mahoney, J. T. (2009). Firm-specific knowledge resources and competitive advantage: The roles of economic- and relationship-based employee governance mechanisms. *Strategic Management Journal*, *30*(12), 1265–1285. https://doi.org/10.1002/smj.787

Wang, H., Zhang, T., Wang, X., & Zheng, J. (2024). Relationship between Occupational Pension, Corporate Social Responsibility (CSR), and Organizational Resilience: A Study on Listed Chinese Companies. *Risks*, *12*(4), 65. https://doi.org/10.3390/risks12040065

Wang, K., Zhang, L., Lei, Z., & Huang, X. (2023). Investigating the impact of digital orientation on economic and environmental performance based on a strategy-structure-performance framework. *International Journal of Logistics Research and Applications*, 1–18. https://doi.org/10.1080/13675567.2023.2215167

Wang, L., & n, J. (2023). Effect of digital transformation on innovation performance in China: Corporate social responsibility as a moderator. *Frontiers in Environmental Science*, *11*, 1215866. https://doi.org/10.3389/fenvs.2023.1215866

Wang, M., Yuan, R., Guan, X., Wang, Z., Zeng, Y., & Liu, T. (2024). The influence of digital platform on the implementation of corporate social responsibility: From the perspective of environmental science development to explore its potential role in public health. *Frontiers in Public Health*, *12*, 1343546. https://doi.org/10.3389/fpubh.2024.1343546

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness–loyalty tradeoff. *Journal of Experimental Social Psychology*, *49*(6), 1027–1033.

Weber-Lewerenz, B. (2021). Corporate digital responsibility (CDR) in construction engineering—Ethical guidelines for the application of digital transformation and artificial intelligence (AI) in user practice. *SN Applied Sciences*, *3*(10), 801. https://doi.org/10.1007/s42452-021-04776-1

Weber-Lewerenz, B., & Traverso, M. (2024). How can corporate digital responsibility CDR be measured in line with construction 4.0. *Sustainable Development*, *33*(1), 287–308.

Wei, S., Wang, L., Jiang, W., Feng, L., & Feng, T. (2023). How eco-control systems enhance carbon performance via low-carbon supply chain collaboration? The moderating role of organizational unlearning. *Corporate Social Responsibility and Environmental Management*, *30*(5), 2536–2554. https://doi.org/10.1002/csr.2501

Wenzel, M. (2005). Misperceptions of social norms about tax compliance: From theory to intervention. *Journal of Economic Psychology*, *26*(6), 862–883. https://doi.org/10.1016/j.joep.2005.02.002

Williams, K. D. (1997). Social ostracism. In *Aversive interpersonal behaviors* (S. 133–170). Springer.

Williams, K. D. (2001). *Ostracism: The power of silence*. Guilford Press.

Williams, R. (2012). Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects. *The Stata Journal: Promoting Communications on Statistics and Stata*, *12*(2), 308–331. https://doi.org/10.1177/1536867X1201200209

Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2023). Corporate Digital Responsibility in Service Firms and Their Ecosystems. In *Journal of Service Research* (Bd. 26, Nummer 2, S. 173–190). https://doi.org/10.1177/10946705221130467

Wirtz, J., & Pitardi, V. (2023). How intelligent automation, service robots, and AI will reshape service products and their delivery. *Italian Journal of Marketing*, *2023*(3), 289–300. https://doi.org/10.1007/s43039-023-00076-1

Wu, W., Shi, J., & Liu, Y. (2024). The impact of corporate social responsibility in technological innovation on sustainable competitive performance. *Humanities and Social Sciences Communications*, *11*(1), 707. https://doi.org/10.1057/s41599-024-03193-0

Wynn, M., & Jones, P. (2023). Corporate Responsibility in the Digital Era. *Information*, *14*(6), 324. https://doi.org/10.3390/info14060324

Xin, D., Yi, Y., & Du, J. (2022). Does digital finance promote corporate social responsibility of pollution-intensive industry? Evidence from Chinese listed companies. *Environmental*

Science and Pollution Research, *29*(56), 85143–85159. https://doi.org/10.1007/s11356-022-21695-9

Xu, J., Yu, Y., Zhang, M., & Zhang, J. Z. (2023). Impacts of digital transformation on eco-innovation and sustainable performance: Evidence from Chinese manufacturing companies. *Journal of Cleaner Production*, *393*, 136278. https://doi.org/10.1016/j.jclepro.2023.136278

Xu, Y., Wang, L., Xiong, Y., Wang, M., & Xie, X. (2023). Does digital transformation foster corporate social responsibility? Evidence from Chinese mining industry. *Journal of Environmental Management*, *344*, 118646. https://doi.org/10.1016/j.jenvman.2023.118646

Yadav, S. S. K., & Mishra, G. (2022). Corporate Digital Responsibility: Perspectives till date and the way ahead. *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, 1474–1479. https://doi.org/10.1109/IC3I56241.2022.10073068

Yadav, S., Samadhiya, A., Kumar, A., Majumdar, A., Garza-Reyes, J. A., & Luthra, S. (2023). Achieving the sustainable development goals through net zero emissions: Innovation-driven strategies for transitioning from incremental to radical lean, green and digital technologies. *Resources, Conservation and Recycling*, *197*, 1–19. https://doi.org/10.1016/j.resconrec.2023.107094

Yang, Q., & Jin, S. (2024). Exploring the Impact of Digital Transformation on Manufacturing Environment, Social Responsibility, and Corporate Governance Performance: The Moderating Role of Top Management Teams. *Sustainability*, *16*(11), 4342. https://doi.org/10.3390/su16114342

Yang, S., Huang, Y., Chan, H.-Y., & Yang, C.-H. (2023). The Impact of Corporate Social Responsibility Practices on Customer Value Co-Creation and Perception in the Digital Context: A Case Study of Taiwan Bank Industry. *Sustainability*, *15*(11), 8567. https://doi.org/10.3390/su15118567

YouGov PLC. (2024). The most popular charities & organisations in the UK: Politics: Yougov ratings. *YouGov PLC*. https://yougov.co.uk/ratings/politics/popularity/charities-organisations/all

Zarte, M., Pechmann, A., & Nunes, I. L. (2022). Knowledge framework for production planning and controlling considering sustainability aspects in smart factories. *Journal of Cleaner Production*, *363*, 1–12. https://doi.org/10.1016/j.jclepro.2022.132283

Zhang, S., Yan, Y., & Yan, X. (2020). A sustainable and preventative warning strategy for the safe production of underground hydrocarbon storage. *Journal of Cleaner Production*, *252*, 119863.

Zheng, L. J., Zhang, J. Z., Kai Ming Au, A., Wang, H., & Yang, Y. (2023a). Leveraging technology-driven applications to promote sustainability in the shipping industry: The impact of digitalization on corporate social responsibility. *Transportation Research Part E: Logistics and Transportation Review*, *176*, 103201. https://doi.org/10.1016/j.tre.2023.103201

Zheng, Y., & Zhang, Q. (2023). Digital transformation, corporate social responsibility and green technology innovation- based on empirical evidence of listed companies in China.

Journal of Cleaner Production, *424*, 138805. https://doi.org/10.1016/j.jclepro.2023.138805

Zhong, X., & Ren, G. (2024). Digitalization and firms' innovation efficiency: Do corporate social responsibility and irresponsibility matter? *The Journal of Technology Transfer*. https://doi.org/10.1007/s10961-024-10075-9

Zhou, Y., Zhang, Q., & Feng, X. (2024). Information Management and User Engagement in Corporate Social Responsibility Communication: A Comparative Analysis of Chinese and Latin American Companies on WeChat and Facebook. *Profesional de La Información*, *33*(3). https://doi.org/10.3145/epi.2024.ene.0317

Zhu, C., Li, N., Ma, J., & Qi, X. (2024). CEOs' digital technology backgrounds and enterprise digital transformation: The mediating effect of R&D investment and corporate social responsibility. *Corporate Social Responsibility and Environmental Management*, *31*(3), 2557–2573. https://doi.org/10.1002/csr.2704

zu Knyphausen-Aufseß, D., & Santarius, T. (2021). Strategic management, the theory of the firm, and digitalization: Reintroducing a normative perspective. *Corporate and Business Strategy Review*, *2*(1), 41–53. https://doi.org/10.22495/cbsrv2i1art4